



Examining Washback on Learning from a Sociocultural Perspective: The Case of a Graded Approach to English Language Testing in Hong Kong

Author: Chi Lai Tsang
University College London

British Council ELT Master's Dissertation Awards 2018: Commendation

Examining Washback on Learning from a Sociocultural Perspective: The Case of a Graded Approach to English Language Testing in Hong Kong

Chi Lai Tsang

A dissertation presented for the degree of
Master of Arts in Applied Linguistics

Supervised by:
Dr Talia Isaacs

University College London, UK
29th August 2017

ABSTRACT

Washback on learning, the effects of testing on various aspects of students' learning (Cheng, Sun & Ma, 2015), has emerged as a focal area of classroom-based language testing research for a quarter of a century. Yet the majority of existing studies have either focused narrowly on observable washback effects (e.g. Latimer, 2009) and/or on discrete mediating factors at the individual test-taker level (i.e. intrinsic factors) in isolation of the wider sociocultural context (e.g. Cheng & Deluca, 2011). To fill this gap in the literature, the present study approaches washback on learning from a sociocultural perspective in the context of the underresearched Hong Kong Diploma of Secondary Education English Examination (HKDSE-English), which, since 2012, has incorporated a graded approach that allows test-takers to choose between an easier or a more difficult task. Using an exploratory sequential mixed-methods design, the goal of the study is threefold: To uncover (1) the washback effects that Hong Kong learners identify following the introduction of the graded approach, (2) learners' self-reported intrinsic and extrinsic mediating factors shaping such washback effects, and (3) the paths of influence between the types of washback effects and the categories of mediating factors. Three sets of semi-structured focus-group interviews involving 12 Hong Kong secondary school learners were first conducted. Following a content analysis, 15 learners' perceived washback effects and 30 items representing learners' self-reported intrinsic and extrinsic mediating variables were inductively generated. Participants' verbatim quotes were then paraphrased and built into a questionnaire, which was administered to another 150 learners. Next, two sets of exploratory factor analyses (EFA) were performed. Results revealed that four major types of washback effects took place not only within but also beyond the classroom, and the mediating variables fell under eight broad categories pertaining to the learners themselves, other stakeholders, and societal influences. Finally, four sets of simultaneous multiple regressions (SMR) were conducted. Results showed that the strongest predictors varied for each washback effect type; yet, they were all significantly affected by at least one intrinsic and one extrinsic factor. Hence, the washback on learning construct appears to be driven by an array of intertwining and potentially competing forces. Overall, the study offers fresh insights into future washback research, including (1) rejecting the notion that washback on learning is rooted only in classrooms, (2) expanding the scope of extrinsic mediating factors to include not only human agents but also societal factors, and (3) reconceptualizing washback on learning as a negotiated construct founded upon the interplay between multiple influences. The study is expected to have considerable impact for educational stakeholders given how it has explicated both the negative washback effects brought about by the graded approach, and possible ways to attenuate such effects.

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION	7
1.1 Research Rationale and Aims	7
1.2 Research Overview	9
CHAPTER 2 LITERATURE REVIEW	11
2.1 Introduction.....	11
2.2 Phase 1 – Studies on the Observable Washback Effects on Learning	11
2.3 Phase 2 – Studies on the Underlying Mediating factors that Shape Learners’ Washback Effects ...	16
2.4 The Gaps that Are to Be Filled in the Area of Washback on Learning	20
2.5 An Introduction to the Graded Approach in HKDSE-English.....	21
2.5.1 English Language Education and Testing in Hong Kong	21
2.5.2 An Overview of HKDSE-English.....	22
2.5.3 The Graded Approach in HKDSE-English	23
CHAPTER 3 METHODOLOGY	25
3.1 Introduction.....	25
3.2 Research Design.....	25
3.3 Participants.....	28
3.4 Research Instruments	30
3.5 Data Collection	31
3.6 Data Analysis	32

3.7 Ethics	34
CHAPTER 4 RESULTS AND DISCUSSION	36
4.1 Introduction.....	36
4.2 Findings.....	36
4.2.1 Research Question 1 – Which washback effects do Hong Kong secondary school learners identify following the introduction of the graded approach in HKDSE-English?	36
4.2.2 Research Question 2 – Which intrinsic and extrinsic mediating factors shape learners’ perceived washback effects?	43
4.2.3 Research Question 3 – What are the relationships between the types of students’ perceived washback effects and the categories of mediating factors, identified and established by the researcher in the present study?.....	51
4.3 Discussion	55
CHAPTER 5 CONCLUSION.....	60
5.1 Introduction.....	60
5.2 Summary of Findings.....	60
5.3 Implications.....	61
5.4 Limitations and Directions for Further Research.....	62

LIST OF APPENDICES

Appendix 1 Research Proposal.....	73
Appendix 2 A Washback Model of Student's Learning (Shih, 2007, p151).....	84
Appendix 3 Interview Guide.....	85
Appendix 4 Washback on Student Learning (WSL) Questionnaire.....	90
Appendix 5 Approach Letter.....	94
Appendix 6 Information Sheet (for Schools).....	95
Appendix 7 Consent Forms (for School).....	98
Appendix 8 Information Sheet (for Parents).....	99
Appendix 9 Consent Form (for Parents).....	101

LIST OF FIGURES AND TABLES

FIGURES

Figure 1 A Diagram of Procedures.....	28
---------------------------------------	----

TABLES

Table 1 The Profiles of the 12 Participants in the Initial Qualitative Phase.....	29
Table 2 Rotated Component Matrix for the First Set of EFA.....	38
Table 3 Rotated Component Matrix for the Second Set of EFA.....	44
Table 4 A Table Illustrating How the Four Sets of SMR Were Run.....	51
Table 5 A Summary of the Results of the First Set of SMR.....	52
Table 6 A Summary of the Results of the Second Set of SMR.....	53
Table 7 A Summary of the Results of the Third Set of SMR.....	54
Table 8 A Summary of the Results of the Fourth Set of SMR.....	55
Table 9 A Summary of the Paths of Influences.....	55

CHAPTER 1

INTRODUCTION

1.1 Research Rationale and Aims

Nearly two decades of research into washback in the field of applied linguistics beginning with Alderson and Wall's (1993) seminal work has given rise to an expanding body of empirical washback literature. Over the years, parallel to the ever-increasing prevalence of large-scale, high-stakes, and often international, testing worldwide (Cheng, 2008; Green, 2006b; Wall & Horák, 2006, 2008) and consequently the growing awareness of the enormous power testing has on the stakeholders (Kunnan, 2010; Shohamy, 2007), the notion of washback has been reconceptualized readily and progressively. To crystalize this gradual transformation process which spans across two decades, the conceptualization of washback has shifted from, initially, a strand of validity (Frederiksen & Collins; 1989; Messick, 1996; Morrow, 1986), to a free-standing construct founded upon the intricate relationships between testing, teaching, and learning (Bailey, 1996; Cheng, 2002; Green, 2007a, 2007b), and, increasingly, to a socially-situated construct which transcends the boundary between washback and impact by bringing together a range of stakeholders and factors within and beyond the classroom context (Booth, 2012; Shih, 2007, 2010). In other words, both the conception and scope of washback has evolved over time. The result of this is that washback is now understood broadly as the effects of tests on "micro-levels of language teaching and learning inside the classroom" (Cheng, Sun & Ma, 2015, p438) as well as on "macro-levels of education and society" (Cheng et al, 2015, p438). Thanks to this, washback studies encompassing a diverse range of areas such as validation of test use (e.g. Bachman, 2005; Bachman & Palmer, 1996; Kane, 2006, 2013), impact evaluation (e.g. Saville & Hawkey, 2004; Green, 2006a, 2006b; Wall & Horák, 2006, 2008), and critical language testing (e.g. Kunnan, 2004, 2010; Shohamy, 1998, 2001, 2007), and across a wide range of contexts including Northern America (e.g. Alderson & Hamp-Lyons; Cheng,

Klinger & Zheng, 2007), Asia (e.g. Wall & Alderson, 1993; Watanabe, 1996; Xie & Andrews, 2012), and Middle East (e.g. Shohamy, Donista-Schmidt & Ferman, 1996), have emerged and flourished in the field.

Among the many areas investigated in empirical work on washback stand the impact studies which examine the systematic changes to teaching and/or learning brought about by testing (Cheng et al, 2015). While earlier impact studies focus largely on washback on teaching – the changes in aspects of teaching as a result of testing (Bailey, 1996), and have offered ample evidence of how testing influences aspects of teaching including teachers' attitude (e.g. Turner, 2006; Wang, 2011), content of teaching (e.g. Green, 2006a, 2006b; Qi, 2004, 2005; Read & Hayes, 2003), and teaching methods (e.g. Alderson & Hamp-Lyons; Hayes & Read, 2004; Mickan & Motteram, 2008), in recent years washback on learning – the effects of testing on various aspects of students' learning (Cheng et al. 2015), has also begun to receive attention. For example, studies tapping into the washback effects among learners (e.g. Green, 2006a; Hawkey, 2006) and the mediating factors shaping such washback effects (e.g. Cheng & Deluca, 2011; Xie & Andrews, 2012) have been conducted to explicate how learners' learning is influenced, or even shaped, by testing. Yet, although these studies have advanced the understanding of washback, the majority of them have either focused rather narrowly on the observable washback effects (e.g. Hawkey, 2006; Latimer, 2009) and/or on discrete mediating factors at the individual level in isolation of the wider sociocultural context (e.g. Cheng, Andrews & Yu, 2011; Cheng & Deluca, 2011). In other words, there is, I argue, insufficient attention paid to the fact that washback on learning is not an independent but a socially-situated phenomenon (Booth, 2012; Shih, 2007) under which learners' observable actions are driven by an array of mediating factors not only within but also beyond the individual level. There is, therefore, a need to go further and to investigate washback on learning more thoroughly and holistically.

In response to this need, the present study is an attempt to approach washback on learning from a broader sociocultural perspective. Drawing upon data gathered from three rounds of focus-group interviews with 12 Hong Kong secondary school learners and a questionnaire administered to another 150 learners across all three school bandings in Hong Kong, this exploratory sequential mixed-methods study

seeks to examine the washback effects Hong Kong secondary school students perceive to have following the introduction of the new graded approach in the Hong Kong Diploma of Secondary Education English Language Examination (HKDSE-English). In particular, this study aims at uncovering, first, students' self-reported changes to their English Language learning in response to the new graded approach, second, the full range of underlying personal and societal mediating factors shaping such washback effects, and, third, the paths of influences between the types of washback effects and the categories of mediating factors identified by the researcher in the present study. The graded approach in HKDSE-English implemented in 2012, which in short allows test-takers to choose between an easier section and a more difficult section in the test, is chosen as the subject of interest because of two primary reasons. First, the washback effects of a graded approach to language testing, whether on teaching or on learning, have yet to be empirically investigated (Smart, Drave & Shiu, 2014) in washback literature and are thus deemed to be examined. Second, and more importantly, given the examination-oriented culture (Berry, 2011; Brown et al, 2009), the stress on meritocracy (Morris & Adamson, 2010), and the fact that the HKDSE-English is a high-stakes test (CDC & HKEAA, 2015), the graded approach is likely to impose significant washback effects on Hong Kong students' English Language learning which, in turn, offers a rich case of analysis. Through investigating the issue from a sociocultural perspective, this study, ultimately, aims at deepening the current understanding of washback on learning in the field.

1.2 Research Overview

The journey of how the present study has progressed to achieve its intended purposes is crystalized in the five chapters in this dissertation. Following this introduction is a literature review where the existing body of work on washback on learning is reviewed across two phases, and where the graded approach in HKDSE-English is introduced. Towards the end of the review when the research gap is identified, the three key research questions guiding the study are presented accordingly. Next, having contextualized the study and built the theoretical groundings on which the study is based, I then proceed to the methodology chapter in which the design and procedures of this mixed-methods study is explained.

What follows next is the results and discussion chapter which presents and discusses the findings in relation to the research questions, draws connections with previous studies, and highlights new and unexpected findings. This dissertation ends with a conclusion chapter which summarizes the findings, discusses implications, identifies limitations, and offers directions for further research.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter begins with a historical overview of the developing body of literature on washback on learning across two key phases: an earlier phase from late 1990s to mid-2000s focusing on the observable washback effects among the learners, and a latter phase from late-2000s to the present time centering on the underlying mediating factors that shape learners' washback effects. The review of these two phases, which highlights the focuses as well as the limitations of each period, is then followed by a section where I identify the gaps that are yet to be filled by existing work, and, subsequently, affirms that a sociocultural approach to washback on learning is a means to fill such gaps. After that, I move on to explain the background against which the research gap is addressed in the present study. This includes a quick overview of, first, English Language education and testing in Hong Kong, second, the aims and structure of HKDSE-English, and, third, the graded approach in HKDSE-English. At the end of the chapter, the three main research questions guiding the present study are presented accordingly.

2.2 Phase 1 – Studies on the Observable Washback Effects on Learning

As highlighted in the introduction, washback on learning was, and perhaps still is, peripheral to washback on teaching in contemporary washback literature (Cheng et al, 2015; Green, 2006a; Zhan & Andrews, 2014). It is, in fact, not until the mid-1990s when the light is finally shed on learners' learning. Following Alderson & Wall's (1993) call for empirical washback studies researching into the effects testing has on aspects of learners' learning including "how learners learn" (p120), "the rate and sequence of learning" (p120), "the degree and depth of learning" (p120), and the assertion of the value of studying learners' learning as put forth by the several studies underscoring the differences between what is taught

and what is learnt (e.g. Alderson & Hamp-Lyons, 1996; Peacock, 1998; Watanabe, 1992), a new phase of washback studies centering on washback on learning began to emerge in the late 1990s. What is most characteristic of this earlier phase of studies conducted between late 1990s and mid-2000s is that their purposes are very often to differentiate washback on learning from the oft-studied washback on teaching, and thereby to establish washback on learning as a field of enquiry in its own right. In most cases, this is done by identifying the very specific noticeable changes learners make to their learning within and beyond classrooms in response to testing. In other words, the focus is placed precisely and, in most cases solely, on what I call ‘the observable washback effects on learning’.

One of the earliest studies from this phase is Cheng’s (1998) study which examined quantitatively the impact the revised English Language examination in the Hong Kong Certificate of Education Examination had on Hong Kong school students’ English Language learning. Drawing upon questionnaire data gathered from 844 secondary five students, Cheng (1998) concluded that the new integrated and task-based approach adopted in the revised examination changed significantly what students learnt and, to a lesser extent, how they learnt in school English lessons. In particular, students were found to prefer and participate considerably more in learning activities that resembled the assessment task types in the communicative-oriented examination such as group discussion and pair work, while at the same time strategically forewent those, such as practicing grammar items and memorizing vocabulary, that were less relevant to the examination. Cheng’s (1998) study, therefore, asserts that learners’ preference for tested skills over untested skills, and their active engagement in intensive test-specific practices in language classrooms are clear indicators of washback on learning.

Cheng’s (1998) finding that learners learn and drill selectively in class is corroborated by subsequent studies in the field (e.g. Ferman, 2004; Green, 2006a, 2006b; Hawkey, 2006; Qi, 2007; Stoneman, 2006; Xie & Andrews, 2011) which span across diverse temporal and spatial spaces. To name two of such studies in this earlier phase, Hawkey’s (2006) study in Central Europe and Qi (2007)’s study in China both yielded sounded supporting evidence. In Hawkey’s (2006) IELTS impact studies, close

classroom observation revealed that learners in an IELTS-preparation class showed a clear preference for task-based activities and the set of micro-skills which they perceived as relevant to the test. Learners' strong desire for test-specific instruction was, in turn, translated and manifested in the classroom in terms of how this had led teachers, who were indeed willing to employ a range of teaching methods and to teach beyond the test, to fall back to IELTS textbooks and focus instructions solely on IELTS. As a result, these learners, just as they wished, learnt and practiced exclusively the test-related skills throughout the course. Similarly, across the continent in China, Qi (2007) also found that Chinese learners sitting for the writing task in the National Matriculation English Test were highly selective in their learning of writing in school English lessons. Instead of learning to write communicatively across a wide range of contexts which was in fact the test developers' intention, once again students were found to focus narrowly on the micro-aspects of writing such as discrete grammar and vocabulary items which they believed were more relevant to the test, and pay little attention to the broader communicative context of writing. These two studies, together with those from many others, therefore, reaffirm that, within language classrooms, learners' strong desire for test-specific instructions (e.g. Hawkey, 2006; Stoneman, 2006), reluctance to learn untested skills (e.g. Qi, 2007; Green, 2006a, 2006b), and intensive drilling on tested items (e.g. Cheng, 1998; Ferman, 2004) are all major washback effects on learning.

Apart from the classroom context, previous studies (e.g. Alderson & Hamp-Lyons, 1996; Ferman, 2004; Gosa, 2004; Hayes & Read, 2004; Mickan & Motteram, 2009; Tsagari, 2009) also indicate that such washback effects often extend beyond language classrooms to learners' personal space. Ferman's (2004) mixed-methods study in Israel and Gosa's (2004) ethnographic study in Romania, for instance, are among the first to put forth the idea that washback to learning is not confined to classroom contexts but is tightly interwoven in the different facets of test-takers' lives. In Ferman's (2004) study, the Israeli learners were found to undertake four specific learning strategies in response to the introduction of the new English Foreign Language Oral Matriculation Test – “intensive learning to the test” (p199), “memorization of test-related tasks” (p199), “self-learning as a compensation strategy” (p200), and “tutor

employment” (p200). These strategies, among which the first two happened primarily in class and the latter two happened in learners’ personal learning environment, thereby illustrated clearly the fact that washback on learning is prevalent both in the language classrooms and in learners’ personal space. Similarly, Gosa’s (2004) diary study concerning the influences of the reformed English component of the Baccalaureate in Romania also revealed that the Romanian learners practiced mock exam tasks intensively in schools, and, surprisingly, even more so in their personal learning environment whenever their expectations were left unfulfilled in school English lessons. These two studies, therefore, collectively and complementarily underscore, first, the argument that washback on learning ought to be reconceptualized to include not only the changes of learning in the classroom but also the changes of learning in learners’ personal space, and second, the fact that learners’ test preparation in their personal learning environment, which in the two studies includes self-learning (Ferman, 2004), employment of tutor (Ferman, 2004), and intensive practice on mock exam tasks (Gosa, 2004), ought to be placed amid the repertoire of observable washback effects on learning as identified in previous studies.

Advancing the argument put forth by Ferman (2004) and Gosa (2004), a number of subsequent studies conducted across a wide variety of countries such as Australia (e.g. Mickan & Motteram, 2008, 2009), Greece (e.g. Tsagari, 2009), Hong Kong (e.g. Chik & Besser, 2011; Stoneman, 2006), and New Zealand (Hayes, 2003; Hayes & Read, 2004) continues to explicate the washback effects happening in learners’ personal environment. To summarize the findings of these studies, the observable washback effects identified fall roughly under two broad categories: explicit ways of practicing the target skills such as ‘drilling on practice materials’ (Mickan & Motteram, 2008; Tsagari, 2009), ‘enrollment in test-focused tutorial classes’ (Chik & Besser, 2011), and ‘employment of personal tutors’ (Stoneman, 2006), and implicit ways of practicing the target skills such as ‘reading extensively’, ‘going to the movies’, and ‘speaking English with friends and family’ (Alderson & Hamp-Lyons, 1996). Therefore, what these studies have contributed to the area is an expansion of the conceptualization of what constitute washback

on learning – the observable washback effects on learners' learning is no longer a self-contained construct residing in classrooms but is one that overflows to the many aspects of learners' personal lives.

Whereas the aforementioned studies center primarily on the washback effects on learners' action, a small body of work in this phase (e.g. Andrews, Fullilove & Wong, 2002; Green, 2007b; Hirai & Koizumi, 2009; Latimer, 2009; Rao, McPherson, Chand & Khan, 2003; Virkiu, 2009; Stecher, Chun & Barron, 2004) goes a step further to examine the washback effects on learners' learning outcome. Although studies into the effectiveness of test-preparation courses have yet to reach consensus (e.g. Green, 2007b; Perrone, 2010; Rao et al, 2003), what appears to be agreed upon in the field is the fact that washback on learning leads to both positive and negative impact on learners' learning outcome, a view that is straightly in line with Bailey (1996) and Alderson & Wall's (1993) hypotheses. On the positive side, as realized in the score gains (e.g. Hirai & Koizumi, 2009; Latimer, 2009), testing brings about improved learning outcome in the set of tested skills given how they are taught, studied, and revised extensively in and out of classrooms. However, on the negative side, testing simultaneously discourages the study of untested skills, and thereby leads to a narrowed set of learning outcome specific to the test (Latimer, 2009; Stecher et al, 2004; Virkiu, 2009). An excellent illustration of this sense of dilemma is offered in Andrews et al's (2002) mixed-methods study. Upon a close analysis of 93 learners' real test performances, Andrews et al (2002) drew the conclusion that the introduction of the oral component to the Hong Kong Advanced-level Use of English Examination influenced Hong Kong students' learning outcome both positively and negatively. While there were general improvements in students' spoken performance, signs of rote-learning of test-related strategies and deterioration of untested skills were also spotted. Hence, Andrews et al's (2002) study illustrates aptly the fact that testing influences not only what learners learn, how learners learn, but also what learners achieve at the very end.

To sum up, what this earlier phase of research has established is the notion that washback on learning is a unique, complex, and multi-faceted construct which is fundamentally different from the oft-studied washback on teaching given how it is driven precisely by learners who respond strategically to

testing, and how its influences stretch way beyond classrooms contexts to learners' personal learning environment. Washback on learning is, therefore, a legitimate area of research in its own right. Yet, while acknowledging their contributions, what should be noted is that the majority of these studies (e.g. Andrew et al, 2002; Cheng, 1998; Gosa, 2004; Hawkey, 2006) focuses only on describing the observable aspects of washback (i.e. learners' actions and learners' learning outcomes). In other words, these studies are by nature rather insensitive to, first, the reasons and factors underneath such observable actions and outcome, and, more importantly, the fact that washback on learning is not homogeneous among learners (Bailey, 1996; Cheng et al, 2015). It is, indeed, a growing sense of awareness towards these inadequacies that drives and powers the second phase of studies into washback on learning.

2.3 Phase 2 – Studies on the Underlying Mediating factors that Shape Learners' Washback Effects

Realizing the inadequacy of not going beyond the observable aspects, the subsequent phase of research from late 2000s onwards, guided by the assumption that washback on learning is different across learners, has shifted from a predominately descriptive approach to an exploratory approach which scrutinizes the sea of mediating factors that shape learners' differential and observable washback effects. In very general terms, these mediating factors fall roughly under two broad categories: 1) intrinsic factors – factors that are internal to the individual learner such as the self-perceived language proficiency (e.g. Fox & Cheng, 2007), intended test use (e.g. Xie & Andrews, 2012), and past experiences (e.g. Stoneman, 2006), and 2) extrinsic factors – factors that are external to the individual learner such as the learning environment (e.g. Zhan & Andrews, 2014), parental expectations (Cheng, Andrews & Yu, 2011), and societal influences (e.g. Chik & Besser, 2011). Moving progressively from the personal to the societal level, the following paragraphs highlight and explain the key factors within the two clusters.

The earliest and indeed the most apparent intrinsic mediating factors recognized in literature is learners' self-perceived language proficiency. Ever since studies into washback on learning began to plunge into the network of mediating factors underneath the observable washback effects, across a

considerable number of qualitative (Fox & Cheng, 2007; Zhan & Andrews, 2014) and quantitative studies (e.g. Cheng et al, 2011; Xie & Andrews, 2012), learners' self-perceived language proficiency has been identified repeatedly as a prominent and decisive factor which profoundly shapes learners' washback effects. For example, Fox & Cheng's (2007) qualitative study which examined the preparatory work ESL students and their native peers did for the Ontario Secondary School Literacy Test revealed that ESL students who believed that their English language proficiency was inferior to their L1 peers engaged considerably more in test preparation activities. Whereas the L1 students did nothing more than reading sample test materials and attending information sessions regarding the test, the ESL students were found to demand for and participate in intensive test preparation in class, and use commercially developed test preparation books within and beyond classrooms over extended period of time. Fox and Cheng's (2007) finding that learners' self-perceived language proficiency led to differential washback effects is substantiated in Cheng et al's (2011) quantitative study which is based in another context where, unlike Ontario, English is a second language to the vast majority of learners. Conducted within the context of the English Language School-based Assessment (SBA) in Hong Kong, Cheng et al (2011) also found that Hong Kong students with higher self-rated English proficiency engaged in significantly more SBA-related activities inside and outside the classroom. Therefore, the studies by Cheng et al (2011) and Fox and Cheng (2007) which are amid the many others (e.g. Stoneman, 2006; Xie & Andrews, 2012), collectively reaffirm the fact that learners' self-perceived language proficiency directs not only what preparatory work learners do (i.e. the actions) but also how hard they practice (i.e. the intensity).

Following the assertion of learners' self-perceived language proficiency as a decisive mediating factor, several studies (e.g. Cheng & Deluca, 2011; Stoneman, 2006; Xie & Andrews, 2012) continue to explicate the range of intrinsic mediating factors shaping learners' observable washback effects. The large-scale quantitative study by Xie & Andrews (2011) in China involving 870 Chinese learners is a classic example clearly representative of this body of work. Drawing insights from Expectancy-value motivation theory, Xie and Andrews (2012) proposed a College Entrance Test (CET) washback model

which attempted to “explain the paths of influences from perceptions of test design and uses to test preparation” (p49). In this model, it was hypothesized that learners’ perception of test design and their intended test use influenced their test preparation both directly and via two motivation factors: value (i.e. perceived stakes of the test) and expectancy (i.e. self-expectation of test success) accordingly. The results of the study suggested that data fitted Xie and Andrews’s (2012) hypothesized model well – the five categories of washback effects including “memorizing vocabulary and model essays, drilling target language skills, and rehearsing test-taking skills” (p59) were all significantly influenced by the four hypothetical intrinsic factors: intended test use, perceptions of test design, perceived stakes of the test, and expectation of test success. Therefore, what Xie and Andrews’s (2012) study has contributed is the addition of the four new intrinsic factors, which are all supported by later studies (e.g. Booth, 2012; Zhan & Andrews, 2014) to the scene where an increasing number of intrinsic factors such as learners’ previous experience (e.g. Cheng & DeLuca, 2011; Stoneman, 2006) has gradually come into play.

Meanwhile, a few researchers (e.g. Chik & Besser, 2011; Green, 2006a) are increasingly aware of the existence of extrinsic mediating factors which are equally powerful in shaping learners’ washback. Green (2006a), for example, argued in his Chinese IELTS impact study that teachers and courses were tremendously influential in shaping the washback effects on learners. As revealed by the data gathered from multiple rounds of questionnaire administered to groups of Chinese learners and their teachers, the narrow test preparation strategies that learners practiced intensively within and beyond the courses were found not to be driven primarily by the learners’ expectations but by the teachers’ expectations. In other words, in Green’s (2006a) study the teachers who possessed the power to decide what ought to be taught and learnt were in fact the ones to direct and drive learners towards such test-focused preparatory work. Whereas Green (2006a) focused essentially on extrinsic factors at the school-level, Chik and Besser (2011) looked more broadly. In their study, Chik and Besser (2011) underscored and explained the influences of schools, private learning centres, and the media in their discussion of how Hong Kong young learners’ learning were shaped by the interaction of such extrinsic factors under the prevalence of international

testing. These two studies, thus, represent very neatly the growing recognition of the power of extrinsic factors in shaping the washback effects on learning.

Interestingly, instances in which intrinsic and extrinsic factors are studied together are scarce. To my knowledge, the work by Booth (2012), Shih (2007, 2009, 2010), and Zhan and Andrews (2014) are, if not the only, amid the very few studies that do so. Among the three qualitative ethnographic studies, Zhan and Andrews's (2014) study concerning the impact of CET in China offers the simplest account of how intrinsic and extrinsic factors worked together to shape learners' washback effects. In their study, Zhan and Andrews (2014) identified four mediating factors that were prevalent among the Chinese learners: "beliefs regarding the CET" (p81), "self-knowledge" (p82), "others' experiences in taking the CET" (p83) and "learning environment" (p84), among which the first two factors were intrinsic and the latter two were extrinsic. These four factors were, then, interwoven together to form a unique "possible CET-self" (p84) which, according to Zhan and Andrews (2014), decided what the resulting washback effects were and to what extent such effects happened, and thereby accounted for the differential washback effects across learners. Compared to Zhan and Andrews (2014), Shih (2007, 2009, 2010) and Booth (2012) presented more developed and systematic models of washback on learning. In Shih's (2007, 2009, 2010) series of ethnographic studies conducted in two Taiwanese university departments, the five washback effects on Taiwanese students' learning including the changes in the 'content of learning' and 'learning strategies' were found to be shaped profoundly by three broad categories of mediating factors under which lists of sub-factors within each category were presented: test factors (12 sub-factors), intrinsic factors (3 sub-factors) and extrinsic factors (16 sub-factors). Shih's (2007, 2009, 2010) finding is, interestingly, mirrored in Booth's (2012) study in Korea which reaffirmed that learners' washback effects were rooted in the "test-taker, community & test complex" (p292) where factors relating to the "test-taker, extended community, and test" (p292) were contained. Therefore, as a whole these three studies, although on a small scale and are primarily qualitative, contribute to the developing field given how they have explicated a handy list of intrinsic and extrinsic factors, and, more importantly, offered a much-needed

perspective into how washback on learning can be studied under a broader socially-situated approach the brings together the intertwining networks of intrinsic and extrinsic factors which are all quintessential in shaping learners' observable washback effects.

To sum up the current phase of research into washback on learning, the developing body of research has gone beyond the mere description of the observable aspects to a close examination of the intrinsic and extrinsic mediating factors influencing learners' washback. Yet, this is not to say the work is comprehensive enough to account for the complexity of the phenomenon. In particular, given how studies on intrinsic factors outnumber that of extrinsic factors, mediating factors beyond the individual level have yet to receive sufficient attention. Furthermore, the fact that most factors, especially the intrinsic ones, are very often studied independently in isolation of the wider socio-cultural context (e.g. Xie & Andrews, 2012), and that research into intrinsic factors and extrinsic factors appear to go parallel to one another with the exception of a few studies (Booth, 2012; Shih, 2007, 2009, 2010; Zhan & Andrews, 2014), indicate that washback has still yet to be considered as a socially-situated phenomenon, a view that is increasingly advocated in the wider fields of applied linguistics (Cheng et al, 2015).

2.4 The Gaps that Are to Be Filled in the Area of Washback on Learning

Following the review of the two phases of contemporary research into washback on learning, it should now become clear that washback on learning is still a developing area in which the many gaps between the three distinctive and often discrete founding blocks: the observable washback effects, intrinsic mediating factors, and extrinsic mediating factors, are left to be filled. Among the many gaps, there are, I argue, two crucial ones that are deemed to be addressed. First, as highlighted in the above discussion, there is a notable lack of studies researching simultaneously into both the intrinsic and extrinsic mediating factors which collectively shape the observable washback effects. Even if an extremely small number of such studies do exist (Booth, 2012; Shih, 2007, 2009, 2010; Zhan & Andrews, 2014), they are all purely qualitative in nature and are thus relatively small scale. Hence, studies adopting

a broader sociocultural approach and incorporating a quantitative component with a larger sample size are clearly and urgently needed. Second, and more importantly, the question of precisely how the categories of intrinsic and extrinsic mediating factors (e.g. personal, school, and societal) relate to the types of observable washback effects (e.g. intensive drilling on tested items, reluctance to learn about the untested items, enrollment in tutorial classes) is left largely unanswered in existing literature. Therefore, quantitative enquiries into the paths of influences from the different categories of mediating factors to the different types of observable washback effects ought to be pursued to offer statistically-sound answers to the question. With these two research gaps in mind, I now proceed to explain the overarching context within which the present study that aims at addressing these two gaps is based.

2.5 An Introduction to the Graded Approach in HKDSE-English

2.5.1 English Language Education and Testing in Hong Kong

Over the past century the intertwining political, historical, economic, societal, cultural, and, not until the recent decade, neoliberal forces (Forestier, Adamson, Han & Morris, 2016) have made the English Language a most valuable asset in Hong Kong. Parallel to the development of the city: from a rural fishing village to a British colony; from a small British port to a regional trading centre; and from a prosperous British city to a global financial hub under the sovereignty of the People's Republic of China, the importance of English has constantly been on the rise (Chan, 2014; Kan & Adamson, 2010; Poon, 2013). In the present day Hong Kong, English is one of the two official languages with statutory power and is the language used widely across a full range of formal and informal contexts such as everyday conversations, classroom teaching and business transactions (Berry, 2011; Bolton, 2011; Evans, 2010).

Given the importance of the English language, it then follows logically that English Language Education has always been a core component of formal schooling in Hong Kong. As stipulated in all curriculum documents released by the Curriculum Development Council (CDC), English Language is one of the eight “key learning areas” (CDC, 2017, p4) that every learner has to pursue throughout their 12

years of compulsory primary and secondary education (CDC, 2004; 2017). Whereas the English Language remains largely as a mandatory core subject at primary schools, English at secondary schools, in addition to being a compulsory subject, is very often also selected as the medium of instruction whenever the school proves itself to be able to pass the threshold set by the CDC (Chan, 2014; Kan & Adamson, 2010). The English Language is, therefore, in every facet of formal schooling in Hong Kong.

To assess learners' achievement in their English Language learning, a vigorous mechanism of English Language testing has long been in place in the Hong Kong schooling system (CDC, 2004, 2017) since the commencement of mass education in the early 1970s. Crafted and implemented by the CDC and the Hong Kong Examinations and Assessment Authority (HKEAA), this mechanism encompasses both summative and formative; high-stakes and low-stakes; centralized and classroom-based English Language assessment (CDC & HKEAA, 2015). As proven in previous empirical studies (e.g. Berry, 2011; Brown et al, 2009), at primary and secondary schools learners are assessed by both low-stakes formative tests in classrooms on a day-to-day basis, and summative final examinations at the end of each school term. Upon learners' completion of the 12 years of compulsory schooling, the HKDSE-English, a high-stakes centralized achievement/admission test, is administered. This test, which is to be explained in the following sections, is a most powerful English Language test which all local learners, teachers, as well as parents and tutorial schools, prepare for extensively over the years (Morris & Adamson, 2010).

2.5.2 An Overview of HKDSE-English

The Hong Kong Diploma of Secondary Education Examination (HKDSE), first introduced in 2012 as a replacement of the two former sets of senior secondary public examination, is a battery of large-scale, high-stakes, standards-referenced tests administered annually in Hong Kong to over 60000 (HKEAA, 2017a) local senior secondary students in their final year of study (i.e. Form 6; Grade 12). Placed amid the three broad categories of subjects offered in HKDSE, HKDSE-English within which the

graded approach is situated is a compulsory test that all test-takers have to take. HKDSE-English, same as the other tests offered in HKDSE, is developed, administered, scored and reported by the HKEAA.

The overall aim of HKDSE-English is “to evaluate candidates’ achievement of the learning targets and objectives of the (English Language) curriculum” (HKEAA, 2016, p1). In other words, the test is at heart an achievement test which sets out to measure the extent to which senior secondary students’ have achieved the learning objectives of the New Senior Secondary (NSS) English Language curriculum as laid down by the CDC and the HKEAA (see CDC & HKEAA, 2015; HKEAA, 2016). Yet, in reality this intended purpose is very often overshadowed by the test’s power to select students for tertiary education at local and overseas institutions. For example, in order to apply for the undergraduate programmes offered by the eight University Grants Committee-funded institutions, students are required to attain a minimum level of Level 3 on the 5-level reporting scale, with Level 5 being the highest, in HKDSE-English (HKEAA, 2017b). The same holds true for the 250 overseas tertiary institutions which recognize the test (HKEAA, 2017c). Therefore, deviating from the intentions, the purpose and use of HKDSE-English, and indeed the whole HKDSE test battery in general, are often to inform universities the proficiency levels of the test-takers so that admission decisions can be made accordingly.

Given that the test sets out to be an achievement test, HKDSE-English is developed around the NSS English Language curriculum guide and is therefore composed of four papers: reading, writing, listening & integrated skills, and speaking, and a school-based component which each carries different weighting (see HKEAA, 2016). In the reading paper and the integrated skills paper which carry 20% and 30% of the subject mark respectively, a graded approach to language testing is adopted. This unique design feature in the world of testing (Smart et al, 2014) is explained further in the section below.

2.5.3 The Graded Approach in HKDSE-English

The graded approach, according to the HKDSE-English test developers who seek to cater for the wide range of candidate abilities previously covered by the two former senior secondary English

Language examinations (Smart et al, 2014), refers to a unique feature of test design in HKDSE-English under which “compulsory and optional sections are administered in the same exam paper sitting” (p258). Under this graded approach, HKDSE-English test-takers in the reading paper and in the listening & integrated skills paper are now granted the freedom to choose either Part B1 – the easier section, or Part B2 – the more difficult section, after they have completed the compulsory Part A. Test-takers attempting Part A and B2 will have access to the full range of possible levels, while the highest attainable level for those attempting A and Part B1 is Level 4 (out of 5). As argued by the test-developers, such a design allows learners the opportunity to select an examination section that is appropriate to their own abilities, and therefore, helps learners of varying language abilities to perform to their best.

While acknowledging the potential this graded approach has in catering for learners’ diverse language abilities and in helping learners to perform to their best, to me what appears to be more interesting is how and to what extent this approach, which by its very nature encourages learners to focus selectively only on either the easier or the more difficult section, influences Hong Kong senior secondary students’ learning. In other words, this new graded approach which is situated within the dynamic Hong Kong context and has yet to be studied in literature is a wonderful window through which I could study, first, the list of learners’ perceived washback effects following the introduction of the graded approach and the full range of intrinsic and extrinsic mediating factors shaping such effects, and second, the relationships between the types of washback effects and the categories of mediating factors uncovered in the present study – the two major research gaps underscored in section 2.4. Adopting a sociocultural perspective, I formulate the following three main research questions which guide the present study:

1. Which washback effects do Hong Kong secondary school learners identify following the introduction of the graded approach in HKDSE-English?
2. Which intrinsic and extrinsic mediating factors shape learners’ perceived washback effects?
3. What are the relationships between the types of students’ perceived washback effects and the categories of mediating factors, identified and established by the researcher in the present study?

CHAPTER 3

METHODOLOGY

3.1 Introduction

Following the literature review which lays the theoretical underpinnings and contextualizes the present study is the methodology chapter which explains the exact procedures, as well as justifications, under which the three guiding research questions were pursued. This methodology chapter begins with an overview of the research design in which the selection of an exploratory sequential mixed-methods design is justified and the two phases of the study (QUAL→QUAN) is introduced. Next, the specifics of the study including the participants, research instruments, data collection, and data analysis involved in the two phases are discussed and justified more extensively in the corresponding sub-section. At the end of the chapter, the measures and procedures undertaken in the study to ensure the ethics which safeguard the rights of the participants are abided by at all stages of the study are presented accordingly.

3.2 Research Design

As neatly summarized by Watanabe (2004) in his landmark methodology paper, washback is “a highly complex rather than a monolithic phenomenon” (p19) and thus ought to be researched with the appropriate method(s). While historically the research paradigm in the field of washback has been oriented towards the qualitative approach (Cheng et al, 2015) which, as Watanabe (2004) argues in his immensely influential paper, takes account of the whole context and the intricate interrelationships between causes and consequences, in recent years the few quantitative washback studies (e.g. Cheng et al, 2012, Xie & Andrews, 2011) have proven the quantitative approach to be equally contributive. To be more precise, the quantitative approach which draws on a larger sample size, despite lacking the power to elicit rich details in naturalistic settings (Brown, 2014; Dörnyei, 2007), is found to possess the unique

power to test hypotheses (e.g. the 14 washback put forth by Alderson and Wall), draw correlations (e.g. between groups of stakeholders in the washback cycle; see Cheng et al, 2002) and construct models (e.g. models of washback; see Xie & Andrews, 2011). In other words, the quantitative approach offers new insights into the research of washback. Therefore, taking the best of both worlds – the depth and the exploratory power offered by the qualitative approach (Brown, 2014; Creswell, 2013; Merriam & Tisdell, 2015), and the breath and the power of quantification offered by the quantitative approach (Dörnyei, 2007; Schreiber & Asner-Self, 2010), in the present study I decided to adopt a mixed-methods design which “gathers both quantitative data and qualitative data, integrates the two, and then draws interpretations based on the combined strengths of both sets of data” (Creswell, 2015, p2).

Among the many mixed-methods designs, the chosen type of design under which the three research questions were pursued in the present study is the exploratory sequential design which studies a phenomenon by “collecting and analyzing quantitative and then qualitative data in two consecutive phases within one study” (Ivankova, Creswell & Stick, 2006, p3). The exploratory sequential design (QUAL→QUAN) is selected to investigate Hong Kong learners’ perceived washback effects following the introduction of the brand-new graded approach in HKDSE-English because of the two special qualities inherent to this design. First, the exploratory power offered in the initial qualitative phase (Creswell, 2013, 2015) allows a close and simultaneous examination of both the list of learners’ perceived washback effects and the array of underlying mediating factors, which, as argued in the literature review, is under-researched in the field. Second, the subsequent quantitative phase makes it possible to draw numerical and statistical correlations between the types of washback effects and the categories of mediating factors uncovered and identified by the researcher in the qualitative phase – a gap that has yet to be addressed in existing literature. The exploratory sequential design is, therefore, an appropriate design under which the two research gaps are addressed and the three research questions are pursued.

Adopting the typical exploratory sequential design, the present study has two main phases as outlined in Figure 1 below. From a helicopter view, the study commences with the qualitative Phase 1

which elicited both the learners' perceived washback effects and their self-reported mediating factors through the conduction of three sets of semi-structured focus-group interview (N=4 in each group). Following a careful content analysis performed on the interview data, two comprehensive lists – one on learners' perceived washback effects and one on their self-reported mediating factors, under which the corresponding quotes from the participants were grouped, were generated. These two lists which offer qualitative answers to the first two research questions were, then, paraphrased into statements and built into a questionnaire which was administered to another 150 learners in the quantitative Phase 2. After this point, the quantitative interview data entered two rounds of data analysis in which further answers to the three research questions were sought. First, two sets of exploratory factor analysis (EFA) were performed to a) the items concerning the washback effects and b) the items concerning the mediating factors respectively, so that the major types of washback effects and the broad categories of mediating factors were identified. Together with the qualitative data in Phase 1, the results of the two sets of EFA offered complete answers to research questions one and two. Next, multiple sets of simultaneous multiple regression (SMR) between the categories of mediating factors (i.e. the independent variables) and each type of washback effects (i.e. the dependent variable) identified through the previous sets of EFA were conducted to examine the paths of influences from the categories of mediating factors to each type of washback effects, which is, precisely, the answer to the final research question. Therefore, as shown above, the present study is designed meticulously in a way that answers to the research questions were sought sequentially and logically. In the following sub-sections, the specifics of the present study which has yet to be discussed are explained in greater details.

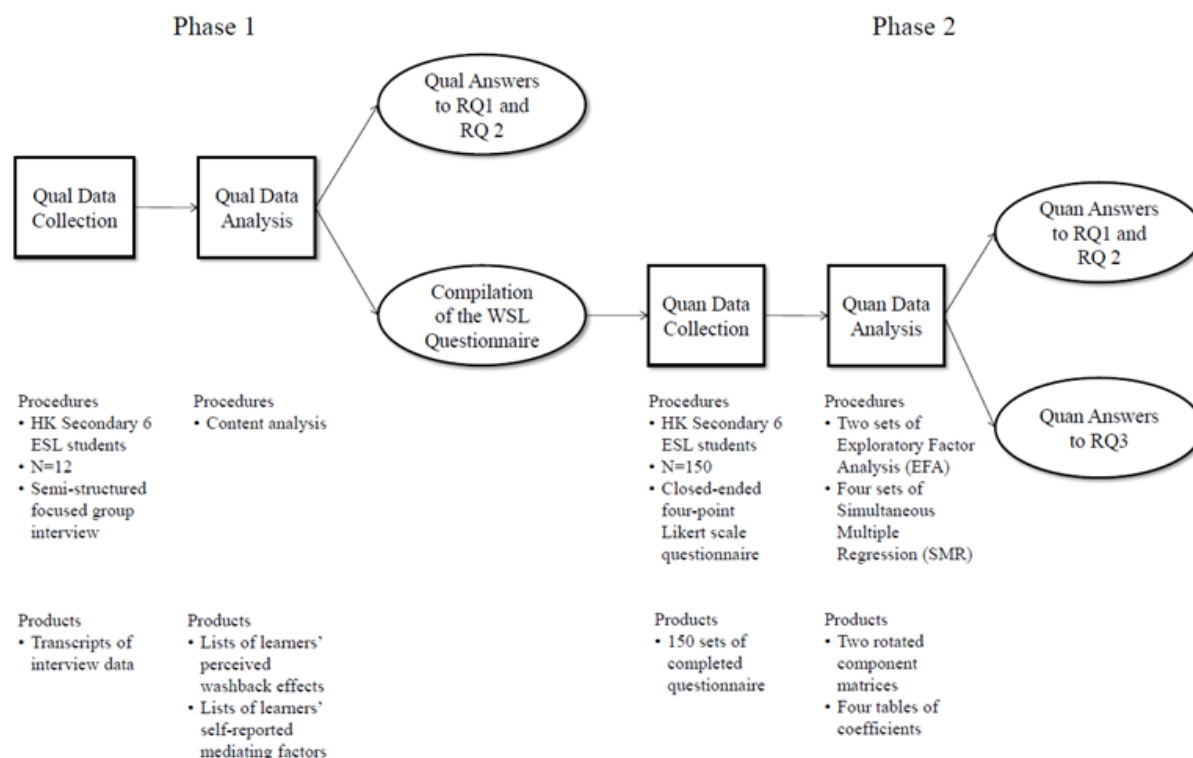


Figure 1 – A Diagram of Procedures

3.3 Participants

Under the overarching exploratory sequential design, participants in the present study were selected on the basis of sequential mixed-methods sampling which is characterized by the employment of first a qualitative sampling strategy in the initial qualitative phase, and then a quantitative sampling strategy in the following quantitative phase (Brown, 2014). In the initial exploratory qualitative phase, 12 Hong Kong secondary six (i.e. Grade 12) ESL students (see Table 1 for their profiles) were selected using maximum variation sampling. Maximum variation sampling, which captures “the widely varying instances of the phenomenon” (Merriam & Tisdell, 2015, p79), was employed in this phase in the sense that, on a macro-level, these 12 students were selected purposefully from three different local secondary schools (i.e. four students from each school) which spanned across and corresponded to the three school bandings in Hong Kong (i.e. Band 1, Band 2, and Band 3; with Band 1 having the highest academic

abilities and Band 3 having the lowest academic abilities). Furthermore, on a micro-level, the four students chosen in each of the three schools, judging from their English Language subject scores in the school examination, were of varying English Language proficiency: one top-scoring student, one high-scoring student, one average-scoring student, and one low-scoring student. Therefore, the macro- and micro-level sampling decisions based on principles of maximum variation sampling worked together to tap into “the whole range of markedly different forms of experience” (Dörnyei, 2007, p128) in the Hong Kong context, and thereby maximized the exploratory power of this qualitative phase in which students’ perceived washback effects and their self-reported mediating factors were elicited.

Participant	School	Age	Academic Abilities	English Subject Score in School Examination
S1	Band-1 English as the Medium of Instruction (EMI) Boy’s School	18	High	174/200
S2		18	High	165/200
S3		18	High	133/200
S4		18	High	101/200
S5	Band-2 Chinese as the Medium of Instruction (CMI) Co-Educational School	19	Average	161/200
S6		18	Average	132/200
S7		18	Average	118/200
S8		18	Average	89/200
S9	Band-3 Chinese as the Medium of Instruction (CMI) Co-Educational School	19	Low	126/200
S10		18	Low	102/200
S11		20	Low	83/200
S12		18	Low	21/200

Table 1 – The Profiles of the 12 Participants in the Initial Qualitative Phase

In a similar vein, the 150 participants in the following quantitative phase, a number that is statistically sufficient for the subsequent EFA (Fields, 2013; Thompson, 2004) and SMR (Keith, 2014), were selected by quota sampling which allows the compilation of a sample that “meets the representations of the population” (Schreiber & Asner-Self, 2010, p85). Therefore, mirroring the three school bandings which comprise the secondary schooling system in Hong Kong, an equal number of Secondary six students was randomly sampled at the same three schools: 50 students from the Band-1 school, 50 students from the Band-2 school, and 50 students from the Band-3 school. Straightly in line with the previous phase, these 150 participants also varied markedly across school bandings, academic abilities, and English Language proficiency, and thus offer a “stratified representative sample” (Dörnyei, 2007, p98) through which quantitative answers to the three research questions were sought.

3.4 Research Instruments

As noted in the previous sections, two research instruments were developed and administered in the present study. The first instrument, which was employed in the initial qualitative phase to elicit learners' perceived washback effects and their self-reported mediating factors, is the three sets of semi-structured focus-group interviews (N=4 in each group) conducted with, first the four Band-1 participants (i.e. S1 to S4), then the four Band-2 participants (i.e. S5 to S8), and finally the four Band-3 participants (i.e. S9 to S12). These three sets of interviews, which drew on the strengths of both semi-structured interviews and focus-group interviews including the flexibility in sequencing questions and closing logical gaps in data (King & Horrocks, 2010; Patton, 2002), and the ability to initiate in-depth discussions on the target issue (Bell, 2014; Gorden, 1975), were in fact the products of a meticulous process of formulation, piloting, and reformulation. At first, a tentative interview guide was formulated with reference to Shih's (2007) "Washback Model of Student's Learning" (p151) (see Appendix 2) which summarized the broad categories of washback effects (e.g. washback on the content of learning, washback on the time spent on learning) and mediating factors (e.g. intrinsic factors, extrinsic factors) identified in the literature. Next, this tentative guide was then passed on to two rounds of piloting, each with four Hong Kong Secondary five students, in which a considerable number of changes were made to re-appropriate the guide to its intended purposes and context. These changes, in short, fell under four main types: 1) eliminating categories less relevant to Hong Kong students, 2) adding in categories relevant to Hong Kong students, 3) merging similar categories, and 4) refining wordings of the questions to ensure they are "progressive, brief, and simple" (Kvale, 1996, p129). The result of these is, ultimately, a three-part revised exploratory interview guide (see Appendix 3) which elicited progressively students' perceptions towards the graded approach in HKDSE-English, their perceived washback effects (i.e. the actions they did in response to the graded approach), and finally their self-reported underlying mediating factors.

Following the administration of the interview and the qualitative data analysis, the washback effects and mediating factors elicited were rephrased into “concise, balanced, and neutral” (Gillham, 2008, p26) statements which were then used as items in the Washback on Students’ Learning (WSL) Questionnaire – the second instrument in the study. This close-ended Likert-scale questionnaire, which allowed the discovery of the clustering of washback effects and mediating factors, as well as the identification of the paths of influences from these identified categories of mediating factors to each type of washback effect, had also undergone a rigorous validation process. First, the draft WSL questionnaire was pretested with four Hong Kong Secondary five students in four 30-minute sections in which each individual student was asked to think-aloud while completing the questionnaire. By doing so, continuous refinement was made to resolve any ambiguity and to ensure the “clarity, accuracy, and appropriateness of the wording for each item” (Xie, 2015, p60). Next, the revised WSL questionnaire was piloted with another 30 Hong Kong Secondary five students which then prompted another round of refinement including 1) replacing the 5-point Likert-scale with a 4-point scale to prevent responses from centering on the neutral position, 2) paraphrasing syntactically similar items to avoid biased clustering in EFA, and 3) further clarifying ambiguous items. These efforts, at the end, led to a finalized two-part questionnaire in which students were asked to rate their extent of agreement towards the list of elicited washback effects in Part A, and their perceived degree of influence of the list of elicited mediating factors in Part B. A copy of this finalized WSL questionnaire is attached in Appendix 4.

3.5 Data Collection

Data in the present study was collected in two consecutive phases. In Phase 1, rich qualitative data was collected through the three sets of semi-structured focus-group interviews (N=4 in each group) conducted with the 12 participants in late June 2017. Each interview was 45-minute long and was administered in Cantonese, the language preferred by the participants. Soon after the interview data was transcribed, categorized, and built into the Washback on Students’ Learning Questionnaire, the questionnaire was administered to the 150 participants in Phase 2 of the study. With a response rate of

100%, 150 completed questionnaires were collected in early July 2017. In the following, the methods through which the qualitative and quantitative data was analyzed is explained more extensively.

3.6 Data Analysis

The sequential exploratory design and the three guiding research questions in the present study collectively led to the data being analyzed through three methods across two consecutive phases. First, in the initial qualitative phase, the interview data collected through the three rounds of interviews was analyzed by a thorough content analysis through which the “meaningful recurring features of the data” (Gibson, 2015, p303), which were in this case students’ perceived washback effects and their self-reported mediating factors, were carefully identified and extracted. To do this, I the researcher, who adopted an inductive approach under which codes emerged naturally from the data (Silverman, 2013), began by studying the transcripts and formulating initial codes through enlisting the apparent washback effects and mediating factors in the dataset. Once these initial codes had been established, illustrative segments of information (i.e. the quotes) under each code were identified and pulled together across participants and rounds of interviews. As more segments were added, refinement to the initial codes, mostly by renaming the codes and by combining the overlapping ones, were made in a “procedural” (Gibson, 2015, p303) manner. These continuous revisions, then, gradually led to the formulation of two refined lists: one on students’ perceived washback effects and one on their self-reported mediating factors, under which corresponding quotes were grouped. To minimize the researcher’s bias and to ensure the internal reliability of the analysis, the final step of this content analysis was to have an independent researcher to check the validity of these two refined lists. After rounds of discussion with the independent researcher in which the differences in opinion (N=7) were spotted and ultimately resolved, the finalized lists of a) students’ perceived washback effects and b) their self-reported mediating factors were generated to offer qualitative answers to research questions one and two, and to inform the design of the WSL questionnaire which was administered in the following phase.

As for the quantitative data collected with the WSL questionnaire, it was analyzed sequentially by two methods. The first one was EFA which unearths the underlying structure of a large set of variables (Hartas, 2015; Thompson, 2004) through “identifying the clusters of variables” (Fields, 2013, p797) that lie within. Given that the first two research questions required the types of students’ perceived washback effects and the categories of their self-reported mediating factors to be identified, and that such intentions were mirrored in the design of the WSL questionnaire, two separate sets of EFA: one on the washback effects (i.e. the 14 items in Part A of the questionnaire) and one on the mediating factors (i.e. the 30 items in Part B of the questionnaire), were performed using SPSS Statistics 24. In general, these two sets of EFA followed the same procedure. To begin with, the sampling adequacy of the two data sets were confirmed by means of Bartlett’s sphericity test (both $<.05$), and the Kaiser- Meyer-Olkin (KMO) index (both >0.5). As the determinants of the two R-matrices were smaller than 0.00001, multicollinearity was proven not to be a problem. Next, having passed the preliminary analysis, the two data sets were passed on for factor extraction. Using Kaiser’s criterion of retaining factors with eigenvalues greater than 1 which was supported by the scree plots, factors were extracted in the two sets of EFA. To improve the interpretability of the two sets of extracted factors, the orthogonal Varimax rotation method which “maximizes the differences between the squared pattern coefficients on a factor” (Thompson, 2004, p42) were performed on the two data sets. The products of these were, ultimately, two rotated factor matrices on which the corresponding extracted factors (i.e. the types of students’ perceived washback effects, and the categories of their self-reported mediating factors) were listed.

Immediately following the two sets of EFA which offered complete answers to the first two research questions were the multiple sets of SMR which dealt with research question three by “determining the extent of influence between one or more variables (i.e. the categories of students’ self-reported mediating factors) on some outcome (i.e. each type of students’ perceived washback effects)” (Keith, 2014, p76). To achieve this, the categories of mediating factors and the types of washback effects as identified in the two previous sets of EFA were first compiled by item parceling under which the scores

of the items (i.e. the discrete washback effects and mediating factors) within each established category/type were averaged. These composite variables were, then, submitted to multiple sets of SMR in which the categories of mediating factors were used as the independent variables and each type of washback effects was used as the dependent variable. In other words, for each type of washback effect, an independent set of SMR using all the categories of mediating factors as dependent variables was performed. By reading the standardized beta values of the dependent variables and the associated t-tests, the paths of influences from categories of mediating factors to each type of washback effect were identified. Hence, the final research question was answered quantitatively.

3.7 Ethics

As all the participants in the present study were young people, a series of measures were undertaken to safeguard their rights. First, in terms of sampling, participants were recruited on the basis of opt-in sampling. Under this approach, an approach letter (see Appendix 5), an information sheet (see Appendix 6) and a consent form (see Appendix 7), which all together explicate in simple language the aims and procedures of the present study, were first sent to the three target schools. Upon receiving the signed consent forms from these schools which thereby indicated that permission was granted, I the researcher began to approach the potential participants at the schools and distribute to them a set of information sheet (see Appendix 8) and parental consent form (See Appendix 9). It was, in fact, only when the signed parental consent forms and verbal consent from these participants were obtained that the process of data collection began. In this way, the participants, as well as their school and parents, were guaranteed to understand the study well and were recruited on a truly voluntary basis. Second, in terms of confidentiality, all data was kept strictly confidential. The data was analyzed only on the researcher's encrypted personal computer in Hong Kong and I the researcher was the only person who had the access to the data. In addition, full anonymity of the participants was achieved through the use of pseudonyms. Third, in terms of the use of findings, participants were well-informed that although results of the study

might be used in future reports and presentations, full anonymity is guaranteed. These three measures, therefore, worked together to ensure that the present study was truly ethical.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

What follows the methodology chapter which explains the design and the procedures of the present study is the results and discussion chapter in which the findings from the two phases are presented and discussed. This chapter opens with a presentation of the findings associated with each of the three research questions. For the first two research questions, the qualitative and quantitative findings collected from the rounds of interviews and the questionnaire are presented integratively, rather than chronologically, so that these two strands of data illuminate one another. As for research question three which by nature could only be answered quantitatively, the corresponding quantitative results are presented. Next, the findings are summarized to offer direct and concise answers to the three research questions. In the last section of the chapter, such findings are discussed in greater depth and beyond the study context through drawing connections with previous studies, and highlighting the new and unexpected findings which are seldom or have yet to be reported in the literature.

4.2 Findings

4.2.1 Research Question 1 – Which washback effects do Hong Kong secondary school learners identify following the introduction of the graded approach in HKDSE-English?

A total of 15 Hong Kong learners' perceived washback effects following the introduction of the graded approach in HKDSE-English, which allows test-takers to choose between an easier section (i.e. Part B1) and a more difficult section (i.e. Part B2), were elicited and identified by the researcher through a careful content analysis performed on the three sets interview data collected in the qualitative Phase 1. To unearth the structure and clustering underneath, these 15 learners' self-perceived washback effects

were submitted to an EFA with Varimax rotation in the subsequent Phase 2. At the preliminary analysis, an overall KMO value of .715 and the fact that KMO values for all individual items far exceeded the acceptable limit of .5, both suggested that the sample was adequate for analysis. An initial analysis was, then, run to obtain eigenvalues for each factor in the data. Four factors, which in combination explained 61.579% of the variance, had eigenvalues over Kaiser's criterion of 1 and were therefore retained. Table 2 shows the factor loadings after rotation. Based on the nature of items that clustered on each factor, the four factors under which the 15 washback effects were grouped were interpreted by the researcher as 1) 'informal ways of training the preferred section outside classroom', 2) 'selective attention in English Language learning', 3) 'intensive paper-and-pencil drills on the preferred section', and 4) 'enrollment in section-focused tutorial classes'. In the following, these four factors, which I refer to as the four types of learners' perceived washback effects hereafter, are explained together with the corresponding quotes elicited in Phase 1 to offer complete answers to research question one.

Rotated Component Matrix^a				
	Component			
	1	2	3	4
15. Reading books at a level of difficulty close to that of the preferred section	.833			
13. Close reading of newspaper articles relevant to the preferred section	.811			
14. Memorization of words relevant to the preferred section	.776			
12. Listening to radio program at a level of difficulty close to that of the preferred section	.744			
7. Disregarding the skills specific to the dispreferred section		.875		
6. Reluctance to drill the dispreferred section		.780		
5. Reluctance to learn about the dispreferred section		.638		
8. Drilling on the skills specific to the preferred section		.576		
3. Eagerness towards the teaching of the preferred section in school English lessons		.506		
2. Drilling on the preferred section of past papers in school English lessons			.827	
1. Drilling on the preferred section of mock papers in school English lessons			.797	
4. Eagerness towards teacher's feedback on his/her own performance in the preferred section			.608	
11. Completion of extra exercises beneficial to the preferred section in out-of-class time			.421	.390
10. Continuous enrollment in tutorials classes targeting at the preferred section				.887
9. Selection of tutorial classes targeting at the preferred section				.860
Extraction Method: Principal Component Analysis.				
Rotation Method: Varimax with Kaiser Normalization. ^a				
a. Rotation converged in 5 iterations				

Table 2 – Rotated Component Matrix for the First Set of EFA

The first type of learners' perceived washback effects identified in the present study is 'informal ways of training the preferred section outside classroom' on which four washback effects are loaded. Among these four washback effects, the first one is 'reading books at a level of difficulty close to that of the preferred section' (item 15), which is well-illustrated in the following quotes:

“I read novels which I think are close to the level of difficulty of Part B2 of the reading paper. I pay special attention to the difficult words which may appear in the exam” (S11, Interview 3)

“I read adolescent literature which I think is at a level close to Part B1. The book series I read contains both short stories and short comprehension exercise. In this way I can test whether I could handle passages in Part B1” (S9, Interview 3)

As shown in these two quotes, the act of reading was performed not for enjoyment but for the preparation of their preferred section in HKDSE-English. In other words, books were read selectively, purposefully and strategically precisely because of the graded approach. This finding is, unsurprisingly, echoed in learners’ selection of newspaper articles and radio programs which were also used as informal means of preparing themselves for the preferred section outside classroom:

“I read South China Morning Post because, first, the level of difficulty of the passages is close to Part B2 and, second, the challenging words might appear in Part B2” (S3, Interview 1)

“Listening to RTHK (Radio Television Hong Kong) is part of my preparatory work. I do so because the topics that appear in Part B2 are usually current issues in society and RTHK covers all of these” (S7, Interview 2)

‘Close reading of newspaper articles relevant to the preferred section’ (item 13) and ‘listening to radio programs at a level of difficulty close to that of the preferred section’ (item 12) are, thus, another two washback effects that are of the same type. Whereas the three washback effects above have broader focuses such as increasing the speed of reading/listening and predicting the topic of reading/listening, the last washback effect in this category, ‘memorization of words relevant to the preferred section’ (item 14) centers particularly on collecting words that might appear in the preferred section:

“Sometimes I will be walking down the streets and scanning for difficult words that might appear in Part B2... These words are everywhere... I try to memorize these words” (S11, Interview 3)

As shown above, these four washback effects which happened informally and outside of classrooms all together constitute the first major type of learners’ perceived washback effects.

The second type of washback effects is ‘selective attention in English Language learning’ which is loaded with five washback effects. On the one hand, learners reported that they embraced the teaching, learning, and practicing of their preferred section given how ‘eagerness towards the teaching of the preferred section in school English lessons’ (item 3) and ‘drilling on skills specific to the preferred section’ (item 8), were prevalent among the learners’ accounts:

“In the integrated skills lessons, our class really appreciates the time when our teacher teaches us how to write with the appropriate tone, register, opening and closing, which are all skills that are only relevant to Part B2” (S4, Interview 1)

“The question types that we e work on a lot in reading lessons are those that appear only in Part B1 – multiple-choice questions, ordering questions, and those that we can directly copy the answer from the passage” (S9, Interview 3)

“... the reading skills that are specific to Part B2 are those we love to drill heavily on. In general, these skills include ‘inferring the meaning of unfamiliar vocabulary items’, ‘reading for main ideas’, ‘summarizing’, and ‘inferencing’” (S1, Interview 1)

While the preferred section was taught and learnt enthusiastically, the dispreferred section did not share the same fate. Instead, ‘reluctance to learn about the dispreferred section’ (item 5) and ‘reluctance to drill the dispreferred section’ (item 6) were immediately apparent in the dataset:

“...our teacher once tried to give us an introduction of Part B2 ... However, nobody in the class seems to listen” (S9, Interview 3)

“It is a waste of time to learn, not to say practice, Part B1” (S4, Interview 1, 2’55)

The reluctance to learn about, let alone practice, the skills in the dispreferred section ultimately led to such skills being neglected in their English Language learning completely and purposefully:

“Part B1 is all about fill-in-the-blanks so the only skill involved is direct copying which is in no way relevant in Part B2” (S4, Interview 1)

“Long question which appears in Part B2 is completely neglected simply because of the fact that it does not appear in Part B1... we lack the ability to construct such a long response anyway” (S9, Interview 3)

Hence, learners’ act of ‘disregarding skills specific to the dispreferred section’ (item 7), alongside the four other washback effects, collectively indicated that the English Language was learnt selectively as a result of the graded approach – the second type of washback effects uncovered in the study.

The third type of washback effects building upon the second is ‘intensive paper-and-pencil drills on the preferred section’ which, as revealed by its four constituents, happened within and beyond the classroom. Within the classrooms, learners reported that ‘drilling on the preferred section of mock HKDSE-English papers’ (item 1) and that of past HKDSE-English papers (item 2) was an everyday practice:

“Drilling Part B2 of the mock papers, which are taken from textbooks and tutorial centers, is basically what we are doing in our English lessons” (S5, Interview 2)

“In the English lessons, we drill a lot on the authentic HKDSE-English past papers because they could tell how we are going to perform in the exam... Of course, we practice only Part B1, but never Part B2” (S9, Interview 3)

In most English lessons which had the standardized structure of 1) paper-and-pencil drills, 2) marking of the exercises, and 3) evaluation, learners developed an ‘eagerness towards teacher’s feedback on his/her performance in the preferred section’ (item 4) especially upon the completion of the drills:

“The marking of Part B2 requires professional judgement because there are many subjective questions... We are eager to receive feedback from our teachers concerning how well we did in Part B2” (S2, Interview 1)

Going beyond the classroom context, learners also reported that ‘completion of further paper-and-pencil exercises beneficial to the preferred section in out-of-class time’ (item 11) was equally common:

“My classmates and I often approach our class teacher and ask for more exercises targeting at the skills tested in Part B2” (S10, interview 3)

Therefore, judging from these four washback effects, it should now become clear that intensive paper-and-pencil drilling on the preferred section, mostly in the form of mock papers, past papers, and drills, was a major type of washback effects among the learners.

The last type of washback effects identified in the present study is ‘enrollment in section-focused tutorial classes’ on which two washback effects are loaded. As illustrated by the following quotes, learners chose tutorial classes selectively and continued to be enrolled on such courses when their desire for section-specific tutoring was satisfied:

“All courses in tutorial schools are classified into two streams – an elite stream and a standard stream in which the content, focus, and duration are all different... given our desire to do well in Part B2, we select the elite stream” (S4, Interview 1)

“I will not drop a course when it does teach me skills specific to Part B1” (S12, Interview 3)

In other words, the washback effects ‘selection of tutorial classes targeting at the preferred section’ (item 9) and ‘continuous enrollment in section-focused tutorial classes’ (item 10) comprised the final type of washback effect.

To sum up the answers to research question one, the present study identified a total of 15 Hong Kong learners' perceived washback effects following the introduction of the graded approach in HKDSE-English. These 15 washback effects, as revealed by the first set of EFA, fell under four major types: 1) 'informal ways of training the preferred section outside classroom', 2) 'selective attention in English Language learning', 3) 'intensive paper-and-pencil drills on the preferred section', and 4) 'enrollment in section-focused tutorial classes'. Readers may refer back to Table 2 for a visual illustration of such.

4.2.2 Research Question 2 – Which intrinsic and extrinsic mediating factors shape learners' perceived washback effects?

A total of 30 Hong Kong learners' self-reported mediating factors shaping their perceived washback effects were identified by the researcher through a content analysis performed on the interview data. In order to uncover the structure that lies underneath, these 30 mediating factors were submitted to an EFA with Varimax rotation. At the preliminary analysis, an overall KMO value of .818 and the fact that KMO values for all individual items far exceeded the acceptable limit of .5, suggested that the sample was adequate for analysis. An initial analysis was then run to obtain eigenvalues for each factor in the data. Eight factors, which in combination explained 68.317% of the variance, had eigenvalues over Kaiser's criterion of 1 and were therefore retained. Table 3 shows the factor loadings after rotation. Based on the nature of items that clustered on each factor, the eight factors under which the 30 mediating factors were grouped were interpreted by the researcher as 1) evaluations made by the teacher, 2) experiences in tutorial schools, 3) societal influences, 4) peer influences, 5) language proficiency, 6) family expectations, 7) knowledge about the examination, and 8) interest towards the English Language. In the following, these eight factors, which I refer to as the eight categories of learners' self-reported mediating factors hereafter, are explained together with the corresponding quotes elicited in Phase 1 to offer complete answers to this research question.

Rotated Component Matrix ^a								
	Component							
	1	2	3	4	5	6	7	8
17. Teacher's advice	.756							
16. Teacher's assessment of individual student's English proficiency	.728							
15. Teacher's expectation	.702							
14. Teacher's selection of course focus	.654							
20. School banding	.584							
19. Materials used in class	.568							
18. Comparisons made by the teacher	.537							
22. Tutorial school's advice		.904						
23. Skills learnt in tutorial class		.835						
21. Advertisement posted by tutorial schools		.619						
29. The importance of HKDSE-English in modern HK society			.708					
30. The examination-oriented culture in HK			.689					
28. Reports published by the media			.553					
27. Posts on forums			.551					
1. Ability to resist external influence			.497					
12. Classmate's selection of examination section				.799				
13. Classmate's preparatory work				.769				
11. Classmate's performance in school examination				.644				
7. Performance in school mock examination					.752			
6. Performance in school assessment tasks					.727			
3. Priority of the English subject in relation to other subjects					.550			
8. English proficiency					.506			
25. Sibling's performance in HKDSE-English						.778		
26. Sibling's advice						.646		
24. Parents' expectations						.437		
2. Personal ambitions						-.412		
5. Knowledge about the examination							.832	
4. Perception of the two examination sections							.812	
9. Availability of time for the English subject								.718
10. Interest towards the English language								.703
Extraction Method: Principal Component Analysis.								
Rotation Method: Varimax with Kaiser Normalization. ^a								
a. Rotation converged in 14 iterations.								

Table 3 – Rotated Component Matrix for the Second Set of EFA

The first category of learners' self-reported mediating factors, which is loaded with a total of seven factors, is 'evaluations made by the teacher'. To begin with, as reported by the learners, 'teacher's selection of course focus' (item 14), which was in essence the result of a class-level evaluation made by the teacher in accordance with the 'school banding' (item 20) and was realized in the 'materials used in class' (item 19), strongly influenced the drilling learners did within and beyond classrooms:

“Our teacher preferred teaching us Part B2. Reasons for not teaching Part B1, such as the fact that the highest achievable grade for Part B1 is only level 4, are brought up from time to time... I think we are ‘brainwashed’ to practice only Part B2” (S3, Interview 1)

“In a Band-3 school, the idea that is delivered by the teacher and shared among everyone is to practice Part B1 only and secure a safe pass at the end” (S10, Interview 3, 12’54)

“...we are conditioned to drill Part B1 only as all the materials we have been exposed to in school English lessons such as exercises, mock papers and past papers are all about Part B1 but never Part B2” (S11, Interview 1)

As shown in the quotes, these three mediating factors collectively discouraged and limited learners’ access to the section that was left untaught by the teacher. Furthermore, in addition to evaluations made on a whole-class-level, learners also stressed that the teacher made influential evaluations on a more personal level. In the study, such micro-level teacher’s evaluations, which were also perceived as important mediating factors, included ‘teacher’s assessment of individual student’s English proficiency’ (item 16), ‘teacher expectations’ (item 15), ‘teacher’s advice’ (item 17) and ‘comparisons made by the teacher’ (item 18), which are illustrated by the following quotes:

“My teacher, judging from my results in the assessment tasks, affirms that my English proficiency is not up to the level of Part B2 so I practice Part B1 exclusively and intensively” (S10)

“My teacher expects me to score well in Part B2 so I work on extra exercises on Part B2 in my personal time” (S3, Interview 1)

“A lot of preparatory work I do is shaped around my teacher’s advice... For example, when my teacher tells me that I should read newspapers for Part B2, I follow suit” (S11, Interview 3)

“Sometimes my teacher praises a few students in the class by telling them they are ready for Part B2. In this way, he is drawing a comparison and implying that the other are not ready and that they should practice Part B2 harder – which is exactly what I do” (S7, Interview 2)

Therefore, as shown above, evaluations made by teachers, be they on the class level or the personal-level, influenced what learners did as well as how hard they did it (i.e. washback effects and the extent of washback), and are therefore a powerful category of mediating factors.

Going beyond the classroom, the second category of mediating factors on which three factors are loaded is ‘experiences at tutorial centres’. As highlighted by the learners, their enrollment in section-focused tutorial class and their intensive drilling on the preferred section were the direct results of ‘advertisement posted by tutorial schools’ (item 21), ‘tutorial school’s advice’ (item 22) and ‘skills learnt in tutorial class’ (item 23):

“Advertisements posted by tutorial schools are everywhere – subconsciously we are told to believe that getting 5** can only be done when we enroll in their section-focused tutorial classes” (S2, Interview 1)

“Tutorial schools always shows us how easy Part B2 is if we master all the ‘skills’. In this way, we believe that Part B2 is not that hard and we start to move away from Part B1 and practice only Part B2” (S10, Interview 3)

“It is a waste not to practice the section-specific skills I learnt in tutorial classes” (S1, Interview 1)

Hence, these three mediating factors establish collaboratively the idea that ‘experiences at tutorial centres’ is an influential category of mediating factors in its own right.

Next, the third category of mediating factors which is made up of five mediating factors is ‘societal influences’. As suggested by its name, this category encompasses mediating factors at a societal level including ‘posts on forums’, (item 27) ‘reports published by the media’ (item 28), ‘the importance of

HKDSE-English in modern HK society’ (item 29) and ‘the exam-oriented culture in Hong Kong’ (item 30). Illustrative quotes of these four factors are as follows:

“The posts on the popular Facebook page ‘Secrets of Prestigious Schools’ are all about how Band-1 students prepare for Part B2. In this sense, the page is actually putting forth the ideas that, first, Band 1 students ought to drill only Part B2, and second, Band 2 and Band 3 students should not even touch Part B2” (S3, Interview 1)

“The media is shaping the idea that practicing Part B1 has not future – it is a section that should be avoided unless your English proficiency is really low” (S8, Interview 2)

“English is going to be useful in the Hong Kong community in whatever disciplines. Therefore, I need a high standard so this is why I practiced Part B2 hard” (S7, Interview 2, 4’35)

“As your future career is decided by your exam results, everyone who wants to lead a successful career are forced to practice Part B2 hard so that they can get a decent grade in English which then leads them to university and then a promising career” (S3, Interview 1, 34’30)

As shown in these quotes, these societal mediating factors consciously and subconsciously led to the creation of stereotypes regarding the two examination sections, which consequently promoted learners to learn and practice selectively. Moreover, as noted by the learners, the degree of influence of these societal factors was also determined by their ‘ability to resist external influences’ (Item 1):

“Timid people are easily subjected to external influences” (S3, Interview 1)

“Those who lack confidence follow whatever the mainstream do” (S6, Interview 2)

Therefore, in short, these societal mediating factors compose a category of prominent mediating factors which are particularly strong when the learner is less of an independent thinker.

Following ‘societal influences’, the fourth category of mediating factors is ‘peer influences’ on which three mediating factors: ‘classmates’ performance in school examination’ (item 11), ‘classmates’ selection of examination section’ (item 12), and ‘classmates’ preparatory work’ (item 13) are loaded. As what is revealed in the following quotes, these three mediating factors had considerable impact on the nature as well as the frequency of learners’ preparatory work.

“When a classmate who is worse than me score higher than me in the school exam, I panic and drill more sets of Part B2 paper so that I could overtake him” (S4, Interview 1)

“When all of my classmates who are of similar level are practicing Part B1, logically I dare not to attempt to practice Part B2” (S10, Interview 3)

“I print whatever my classmates print (Part B2)... I attend whatever tutorial classes my classmates attend” (S6, Interview 2)

Thus, as shown above, ‘peer influences’ is an influential category that ought to be taken into account.

The fifth category of mediating factors is ‘language proficiency’ which comprised four mediating factors. Among the four factors, the first two factors ‘performance in classroom assessment tasks’ (item 6) and ‘performance in school mock examination’ (item 7) offered objective assessment of learners’ language proficiency which then informed them what level they were at and what ought to be done:

“Achieving a high score in quizzes consistently affirms that I should go beyond Part B1 and attempt Part B2. Thus, I drill Part B2 intensively” (S11, Interview 3)

“Not doing well in the school mock examination is the final warning that I should practice Part B2 much harder and take those intensive tutorial courses focusing at Part B2” (S10, Interview 3)

These continuous objective assessment tasks gradually constructed a ‘self-perceived language proficiency’ (item 8) which also exerted great influences on their preparatory work:

“As an EMI student who knows the English language, of course I have no intention to even look at Part B1” (S1, Interview 1)

“My limited knowledge about the (English) language leads me to the more achievable Part B1” (S11, Interview 3)

In some cases, learners’ self-perceived language proficiency decided the ‘priority of the English subject in relation to other subjects’ (item 3) which, in turn, powerfully shaped learners’ washback effects:

“English is my strength so it is the subject that I need to get an excellent grade in. Thus, I practice my Part B2 intensively – I drill the papers in class as well as out-of-class; I attend those tutorial classes; I read newspapers and listen to English songs in my personal time” (S1, Interview 1)

As shown above, language proficiency is a truly decisive category of mediating factors.

As for the sixth category, it is ‘family expectations’ under which four mediating factors are grouped. In the following, quotes illustrating how the first three factors: ‘parents’ expectations’ (item 24), ‘sibling’s performance in HKDSE-English’ (item 25), and ‘sibling’s advice’ (item 26), shaped learners’ washback effects are presented:

“My parents expect me to get a pass so I train part B1 intensively so that I can get a safe pass and thus meet their expectations” (S9, Interview 3)

“Comparisons with siblings always matters – every younger brother aims at doing better than his elder brother so I train very hard on Part B2 so that I could surpass him” (S4, Interview 1)

“My brother who is of similar level as me advises me to attend tutorial classes focusing on Part B2 just like he did. Given our similarities, it makes sense for me to follow suit” (S4, Interview 1)

Interestingly, learners’ desire to meet family expectations were also reflected in their ‘personal ambitions’ (item 2), which is, in fact, the last mediating factor in this category.

After ‘family expectations’ comes the seventh category ‘knowledge about the examination’ which comprised two mediating factors: ‘perception of the two examination section’ (item 4) and ‘knowledge about the examination’ (item 5); among which the former concerns the stereotypes learners had for the two exam sections, and the latter concerns how well learners know the examination.

“Part B2 is for those who are proficient in English and have higher aims, whereas Part B1 is for those who struggle with English, lack exposure, and hope for a safe pass. Because of this, I abandon Part B1” (S1, Interview 1)

“Even if you score full marks in Part B1, at most you can only get a level 4. Therefore, practicing Part B2 is essential for those who want to achieve a better score” (S10, Interview 3)

As what was apparent in the quotes, learner’s perceptions of and knowledge about the examination combined to form a category of mediating factors which influenced their preparatory work.

Last but not least is the category ‘interest towards the English Language’ which is made up of ‘personal interest’ (item 10) and ‘availability of time for the English subject’ (item 9). As shown in the following quote, interest towards the language, which is mirrored in the allocation of time for the subject, is a significant determiner of learners’ washback effects.

“As a person who loves the English language, I find it interesting to practice and work on Part B2 tasks which are a lot more authentic and challenging. Therefore, I do preparatory work for my Part B2 not only within but beyond classroom contexts” (S1, Interview 1)

To sum up answers to research question two, 30 Hong Kong learners’ self-reported mediating factors were identified. These 30 factors, as revealed by the second set of EFA, fell under eight categories: 1) evaluations made by the teacher, 2) experiences in tutorial schools, 3) societal influences, 4) peer influences, 5) language proficiency, 6) family expectations, 7) knowledge about the examination, and 8) interest towards the English Language. Readers may refer to Table 3 for a visual illustration.

4.2.3 Research Question 3 – What are the relationships between the types of students’ perceived washback effects and the categories of mediating factors, identified and established by the researcher in the present study?

The answers to research question three were obtained through four sets of SMR in which the eight categories of mediating factors were simultaneously entered as the independent variables across all the four sets, and each of the four types of washback effects was computed as the dependent variable one set after another (see Table 4 for an illustration). In the following, the results of these four sets of SMR are presented to offer quantitative answers to this research question.

	Dependent Variable	Independent Variable
Set 1	The first type of washback effect	The eight categories of mediating factors
Set 2	The second type of washback effect	
Set 3	The third type of washback effect	
Set 4	The fourth type of washback effect	

Table 4 – A Table Illustrating How the Four Sets of SMR Were Run

The first set of SMR used the first type of washback effect ‘informal ways of training the preferred section outside classroom’ as the dependent variable and the eight categories of mediating factors as independent variables. Overall, the eight categories combined to form a model ($R^2=.147$) that is a significant fit of the data overall ($p=.003$). However, as summarized in Table 5, only three out of eight independent variables: ‘family expectations’ (standardized $\beta=.239$; $p=.032$), ‘language proficiency’ (standardized $\beta=-.188$; $p=.046$) and ‘interest towards the English Language’ (standardized $\beta=.181$; $p=.037$) were proven to be statistically significant predictors of the dependent variable. In other words, what the first set of SMR revealed is that three categories of mediating factors: 1) ‘family expectations’, 2) ‘language proficiency’, and 3) ‘interest towards the English Language’, had a direct effect on the first type of washback effect – ‘informal ways of training the preferred section outside classroom’.

Coefficients^a					
	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	1.677	.404		4.154	.000
Evaluations made by teachers	.076	.101	.067	.757	.450
Experiences at tutorial schools	-.056	.088	-.061	-.641	.523
Peer influences	.154	.082	.179	1.878	.062
Societal influences	-.104	.098	-.115	-1.065	.289
Family expectations	.227	.105	.239	2.165	.032
Language Proficiency	-.246	.122	-.188	-2.017	.046
Knowledge about the exam	.091	.082	.100	1.115	.267
Interest towards the Language	.180	.085	.181	2.109	.037

a. Dependent Variable: Informal ways of training the preferred section outside classroom

Table 5 – A Summary of the Results of the First Set of SMR

Employing the same set of independent variables, the second set of SMR used the second type of washback effect ‘selective attention in English Language learning’ as the dependent variable. Overall, the eight independent variables combined to form a model ($R^2=.132$) that is a significant fit of the data overall ($p=.009$). As summarized in Table 6, three out of eight independent variables: ‘evaluations made by the teacher’ (standardized $\beta=.238$; $p=.009$), ‘language proficiency’ (standardized $\beta=-.263$; $p=.006$) and ‘knowledge about the examination’ (standardized $\beta=.216$; $p=.018$) were proven to be statistically significant predictors of the dependent variable. Therefore, the second set of SMR concluded that three categories of mediating factors: 1) ‘evaluations made by the teacher’, 2) ‘language proficiency’, and 3) ‘knowledge about the examination’, are those that had a direct effect on the second type of washback effect – ‘selective attention in English Language learning’.

Coefficients^a					
	Unstandardized Coefficients		Standardized Coefficients		Sig.
	B	Std. Error	Beta	t	
(Constant)	2.007	.347		5.786	.000
Evaluations made by teachers	.231	.087	.238	2.665	.009
Experiences at tutorial schools	.012	.075	.016	.163	.871
Peer influences	.091	.071	.123	1.282	.202
Societal influences	-.065	.084	-.084	-.776	.439
Family expectations	-.063	.090	-.077	-.694	.489
Language Proficiency	-.294	.105	-.263	-2.804	.006
Knowledge about the exam	.168	.070	.216	2.398	.018
Interest towards the Language	.081	.073	.096	1.103	.272

a. Dependent Variable: Selective attention in English Language learning

Table 6 – A Summary of the Results of the Second Set of SMR

In the same vein, the third set of SMR kept the same independent variables and used the third type of washback effect ‘intensive paper-and-pencil drills on the preferred section’ as the dependent variable. Overall, the eight independent variables combined to form a model ($R^2=.109$) that is a significant fit of the data ($p=.035$). However, as summarized in Table 7, among the eight independent variables, ‘evaluations made by the teacher’ (standardized $\beta=.208$; $p=.023$), and ‘knowledge about the examination’ (standardized $\beta=.201$; $p=.030$), were the only two that were statistically significant predictors. Hence, the third set of SMR identified that these two categories of mediating factors: ‘evaluations made by the teacher’, and ‘knowledge about the examination’, had a direct effect on the third type of washback effect – ‘intensive paper-and-pencil drills on the preferred section’.

Coefficients^a					
	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	2.903	.291		9.987	.000
Evaluations made by teachers	.167	.073	.208	2.297	.023
Experiences at tutorial schools	.103	.063	.160	1.634	.105
Peer influences	-.102	.059	-.168	-1.724	.087
Societal influences	.042	.070	.065	.594	.554
Family expectations	-.134	.075	-.200	-1.774	.078
Language Proficiency	-.039	.088	-.042	-.445	.657
Knowledge about the exam	.129	.059	.201	2.196	.030
Interest towards the Language	-.026	.062	-.037	-.415	.679

a. Dependent Variable: Intensive paper-and-pencil drills on the preferred section

Table 7 – A Summary of the Results of the Third Set of SMR

As for the last set of SMR which again retained the same set of independent variables, the dependent variable that was used is the fourth type of washback effect ‘enrollment in section-focused tutorials classes’. Overall, the eight independent variables combined to form a model ($R^2=.205$) that is a significant fit of the data overall ($p=.000$). As summarized in Table 8, three independent variables 1) ‘experiences at tutorial schools’ (standardized $\beta=.218$; $p=.020$), 2) ‘peer influences’ (standardized $\beta=.184$; $p=.047$), and 3) ‘knowledge about the examination’ (standardized $\beta=.259$; $p=.003$), were proven to be statistically significant predictors of the dependent variable. Therefore, what the fourth set of SMR established is the idea that two categories of mediating factors: 1) ‘evaluations made by the teacher’, and 2) ‘knowledge about the examination’, had a direct effect on the fourth type of washback effect – ‘enrollment in section-focused tutorials classes’.

Coefficients^a					
	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	1.500	.483		3.106	.002
Evaluations made by teachers	-.179	.120	-.127	-1.488	.139
Experiences at tutorial schools	.248	.105	.218	2.361	.020
Peer influences	.197	.098	.184	2.003	.047
Societal influences	.085	.117	.075	.727	.469
Family expectations	-.011	.125	-.009	-.084	.933
Language Proficiency	-.049	.146	-.030	-.335	.738
Knowledge about the exam	.293	.098	.259	2.996	.003
Interest towards the Language	-.105	.102	-.086	-1.029	.305

a. Dependent Variable: Enrollment in section-focused tutorials classes

Table 8 – A Summary of the Results of the Fourth Set of SMR

To sum up, the results of the four sets of SMR offer direct quantitative answers to research question three. In short, what was apparent from these sets of SMR is that each type of washback effect was, indeed, driven by its very specific categories of mediating factors. Readers may refer to Table 9 for summary of these paths of influences.

	Informal ways of training the preferred section outside classroom	Selective attention in English Language learning	Intensive paper-and-pencil drills on the preferred section	Enrollment in section-focused tutorials classes
Statistically Significant Variables	family expectations (standardized $\beta=.239$; $p=.032$)	evaluations made by the teacher (standardized $\beta=.238$; $p=.009$)	evaluations made by the teacher (standardized $\beta=.208$; $p=.023$)	experiences at tutorial schools (standardized $\beta=.218$; $p=.020$)
	language proficiency (standardized $\beta=.188$; $p=.046$)	language proficiency (standardized $\beta=.263$; $p=.006$)	knowledge about the examination (standardized $\beta=.201$; $p=.030$)	peer influences [†] (standardized $\beta=.184$; $p=.047$)
	interest towards the English Language (standardized $\beta=.181$; $p=.037$)	knowledge about the examination (standardized $\beta=.216$; $p=.018$)		knowledge about the examination (standardized $\beta=.259$; $p=.003$)

Table 9 – A Summary of the Paths of Influences

4.3 Discussion

The present mixed-methods study which adopts a broad sociocultural approach to the research of washback on learning has yielded three blocks of findings which correspond to the three guiding research

questions: 1) Hong Kong learners' perceived washback effects following the introduction of the graded approach in HKDSE-English, 2) Hong Kong learners' self-reported mediating factors shaping such washback effects, and 3) the paths of influence between the types of washback effects and categories of mediating factors uncovered in the study. In what follows, these three blocks of findings are discussed one after another beyond the present study context in the sense that, first, parallels between findings in the study and that of the wider washback literature are drawn, and, second, new and unexpected findings that are seldom or yet to be discussed in literature are highlighted.

The first block of findings in the present study is the 15 washback effects identified by Hong Kong learners which, as proven by the first set of EFA, fall under four major types: 1) 'informal ways of training the preferred section outside classroom', 2) 'selective attention in English Language learning', 3) 'intensive paper-and-pencil drills on the preferred section', and 4) 'enrollment in section-focused tutorial classes'. First of all, a close look at these four types reveals that the latter three are closely in line with findings in the contemporary washback literature. For example, selective attention in English Language learning and intensive paper-and-pencil drilling on tested items are acknowledged as major washback effects on students' learning in previous studies in a wide range of contexts including Central Europe (e.g. Hawkey, 2006), China (e.g. Cheng, 1998; Green, 2006a; Qi, 2007) and Greece (Tsagari, 2009). Similarly, enrollment in tutorial classes is also highlighted in a number of studies (e.g. Ferman, 2004; Mickan & Motteram, 2008; Stoneman, 2006) as one of the prominent washback effects. The discovery of these three types of washback effects in the present study, therefore, confirms and adds weight to the argument that such types of washback effects are prominent indicators of washback on learning.

Compared to these three types, the first type of washback effect 'informal ways of training the preferred section outside classroom' is far less elaborated in the literature. While occasionally there are studies (e.g. Alderson & Hamp-Lyons, 1996) in which this particular type is superficially named, detailed accounts and explanations of why these informal ways of learning are classified as washback on learning are scarce. The present study, thus, contributes to filling this gap in literature by, first, illustrating in great

details how these informal ways of learning are precisely the specific actions learners take in response to testing, and second, enlisting the washback effects within this under-researched type.

At a more holistic level, these four types, when put together as a whole, demonstrate a progressive transmission of washback effects across varying venues and contexts: from classrooms, to tutorial schools, and finally to learners' personal environment – a finding that is in agreement with the current understanding of washback on learning (e.g. Ferman, 2004; Gosa, 2004; Hayes & Read, 2004; Mickan & Motteram, 2009; Tsagari, 2009). These four types of washback effects, therefore, post further challenges to the oft-criticized idea that washback on learning is rooted in classrooms (e.g. Cheng, 1998, Hawkey, 2006) and powerfully reaffirm the argument that washback on learning penetrates through the many contexts as well as facets of learning surrounding the learners.

Following the first block of results, the second block of results, the 30 Hong Kong learners' self-reported mediating factors, is the next to be discussed. As revealed by the second set of EFA, these 30 factors fall under eight categories: 1) evaluations made by the teacher, 2) experiences in tutorial schools, 3) societal influences, 4) peer influences, 5) language proficiency, 6) family expectations, 7) knowledge about the exam, and 8) interest towards the English Language. To begin with, the three categories of intrinsic factors identified in the present study: language proficiency, knowledge about the examination, and interest towards the English Language, among which the first two are widely recognized (e.g. Fox & Cheng, 2007; Xie & Andrews, 2012) in literature and the latter remains largely unheard, offer new insights into the existing categories of intrinsic factors. In particular, by highlighting how interest in the language has a direct effect on washback, the present study brings this new category to the scene.

Next, the five categories of extrinsic mediating factors identified in the study: 1) evaluations made by the teacher, 2) peer influences, 3) experiences in tutorial schools, 4) family expectations and 5) societal influences, confirmed the existence of a range of under-researched categories of extrinsic mediating factors enlisted in Shih's (2007) washback model. Although some of the extrinsic mediating

factors within these five categories such as ‘teacher’s advice’ and ‘materials used in class’ were already explained in the few previous studies (Booth, 212; Green, 2006a), in the present study a considerable number of extrinsic factors, particularly those at the wider societal level such as reports published by the media and posts on forums, are still left untouched and are yet to be recognized as extrinsic mediating factors in the literature. Therefore, by elaborating on these under-researched categories of extrinsic factors and by uncovering a considerable number of extrinsic mediating factors which have yet to be identified in the existing literature, the present study adds new dimension to the existing categories of extrinsic mediating factors. To be more precise, the present study has illustrated powerfully the argument that extrinsic mediating factors encompass not only human agents in learners’ immediate learning environment but also the wider social realities in which learners are situated.

The last block of findings discussed in this section concerns the paths of influences between each type of washback effect and its corresponding categories of intrinsic and extrinsic mediating factors. In short, these paths of influences indicate three crucial properties of washback on learning. First, given that each type of washback effect is significantly affected by at least one category of intrinsic factors and one category of extrinsic factors, washback on learning is fundamentally a construct driven by an array of intertwining forces (i.e. the intrinsic and extrinsic mediating factors), which is exactly the gist of phase 2 of washback studies (as discussed in Section 2.3). Second, precisely because of the fact that categories of intrinsic and extrinsic mediating factors are always in play behind every type of washback effects, washback on learning is essentially the product of learners’ careful thinking and strategic negotiation between the sea of (sometimes conflicting) mediating factors through which learners’ likelihood of achieving their desired outcome is maximized. In other words, in contrary to the current thinking (e.g. Cheng et al, 2015), washback on learning is less about imposition but more about negotiation. Third, on a more micro-level, the agent(s) within the context in which the corresponding type of washback effect occurs (e.g. tutors and classmate in a section-focused tutorial class), though not always, very often possess the power to influence washback on learning. Therefore, following this line of logic, a way to

manipulate washback on learning is, theoretically, to alter the way these agents treat the learner. On this note, I end the discussion section and proceed to the final chapter in this dissertation.

CHAPTER 5

CONCLUSION

5.1 Introduction

The conclusion chapter in this dissertation marks the end of the long journey of how the intended purposes put forth at the beginning of the study were achieved one step after another over time. In this chapter, a summary of findings in which the findings to the three research questions are summarized is first presented. After that, the implications are discussed. In the end, limitations of the present study and directions for future research are identified and presented accordingly.

5.2 Summary of Findings

This exploratory sequential mixed-methods study (QUAL→QUAN), which draws upon data collected from three sets of semi-structured focus-group interviews (N=4 in each group) with 12 Hong Kong secondary six learners and a questionnaire administered to another 150 Hong Kong learners, is structured in a way that answers to the three guiding research questions are sought in a logical manner. In the following, the answers to these three research questions are summarized and presented.

Frist, regarding research question 1 which aims at uncovering Hong Kong learners' perceived washback effects following the introduction of the graded approach in HKDSE-English, 15 learners' perceived washback effects were identified. These 15 washback effects, as revealed by the first set of EFA, fell under four major types: 1) 'informal ways of training the preferred section outside classroom', 2) 'selective attention in English Language learning', 3) 'intensive paper-and-pencil drills on the preferred section', and 4) 'enrollment in section-focused tutorial classes'.

Next, regarding research question 2 which seeks to identify Hong Kong learners' self-reported mediating factors, 30 learners' self-reported mediating factors were identified. As revealed by the second set of EFA, these 30 factors fell under eight categories: 1) evaluations made by the teacher, 2) experiences in tutorial schools, 3) societal influences, 4) peer influences, 5) language proficiency, 6) family expectations, 7) knowledge about the examination, and 8) interest towards the English Language.

Finally, regarding research question 3 which studies the paths of influences from categories of mediating factors to each type of washback effect, each type of washback effect was found to be driven by its very specific categories of mediating factors. These paths of influences are summarized in Table 9.

As shown above and previously, the present study achieves its intended aims given how complete answers to the research questions were obtained in its qualitative and quantitative phases.

5.3 Implications

The present study is expected to have considerable implications for researchers in the field of washback, and the HKEAA. To begin with, the study contributes to the research on washback on learning given how it has addressed the lack of a) studies researching into both intrinsic and extrinsic mediating factors, and b) quantitative enquiries into the paths of influences between mediating factors and observable washback effects. Furthermore, using the graded approach as the lens through which washback on learning is studied; along the research process this study has come up with new findings (e.g. the expanded lists of intrinsic and extrinsic mediating factors and the paths of influences), as well as a number of sound arguments including 1) rejecting the notion that washback on learning is rooted only in classrooms, 2) expanding the scope of extrinsic mediating factors to include not only human agents but also societal factors, and 3) reconceptualizing washback on learning as a negotiated rather than imposed construct, which are all of great value to the developing field of washback on learning.

In addition, the present study is also particularly useful for the HKEAA as this study is indeed the first to investigate the washback effects initiated by the graded approach despite the fact that it has been

implemented in the high-stakes HKDSE-English for five years. The present study, which identified the four major types of washback effects, therefore, come at the right time to inform HKEAA that the graded approach, in contrary to their intentions, has led to negative washback effects particularly in terms of learners' excessive paper-and-pencil drills and their selective attention in English Language learning. Moreover, results in this study such as the lists of mediating factors and paths of influences can also be used diagnostically in the sense that the causes of each type of negative washback effects are made apparent by the results of SMR. Based on these, redressive actions can be made more easily.

5.4 Limitations and Directions for Further Research

The present study has two major limitations inherent to the design of the study. First, non-probability quota sampling in the quantitative phase, which is a compromise made owing to issues of practicality, reduces the generalizability of the results of the study. Therefore, even if efforts are made to compile a sample as representative as possible, one risks the accuracy of the data if the results are to be generalized to a wider population. Second, the selection of SMR instead of more sophisticated statistical procedures such as structural equation modelling also limited the extent of analysis that could be done. For example, with SMR, the creation of an explanatory model is almost impossible.

In view of these limitations, future research studies are advised to employ more sophisticated statistical procedure such as structural equation modelling for the creation of explanatory models which explain how the categories of mediating factors interact and work together to shape the various types of washback effects. Furthermore, instead of relying solely on learners' self-reported data, researchers are suggested to utilize ethnographic or observational data in their future research on the topic. After all, washback on learning is complex and multi-faceted construct which ought to be research with multiple and carefully selected methods if it is to be understood in more depth.

(16455 words)

Bibliography

- Alderson, J. C. & D. Wall (1993). 'Does washback exist?' *Applied Linguistics*, 14, 115–129.
- Alderson, J. C. & L. Hamp-Lyons (1996). TOEFL preparation courses: A case study. *Language Testing* 13, 280–297.
- Bachman, L. F. (2005). Building and supporting a case for test use. *Language Assessment Quarterly*, 2(1), 1–34.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests* (Vol. 1). Oxford University Press.
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13, 257–279.
- Bell, J. (2014). *Doing Your Research Project: A guide for first-time researchers*. McGraw-Hill Education (UK).
- Berry, R. (2011). Assessment Trends in Hong Kong: Seeking to Establish Formative Assessment in an Examination Culture. *Assessment in Education: Principles, Policy & Practice*, 18(2), 199–211.
- Bolton, K. (2011). Language policy and planning in Hong Kong: Colonial and post-colonial perspectives.
- Booth, D. K. (2012). *Exploring the washback of the TOEIC in South Korea: a sociocultural perspective on student test activity*. Ph.D. dissertation, The University of Auckland.
- Brown, G. L., Kennedy, K. J., Fok, P. K., Chan, J. S., & Yu, W. M. (2009). Assessment for Student Improvement: Understanding Hong Kong Teachers' Conceptions and Practices of Assessment. *Assessment in Education: Principles, Policy & Practice*, 16(3), 347–363.
- Brown, J. (2014). *Mixed methods research for TESOL*. Edinburgh University Press.

- Carless, D. (2004). Issues in teachers' reinterpretation of a task-based innovation in primary schools. *Tesol Quarterly*, 639-662.
- Curriculum Development Council. & Hong Kong Examinations and Assessment Authority. (2015). *English Language Curriculum and Assessment Guide (Secondary 4 - 6)*. Retrieved from http://www.edb.gov.hk/attachment/en/curriculum-development/kla/eng-edu/Curriculum%20Document/EngLangCAGuide_Nov15.pdf
- Chan, J. Y. (2014). Fine-tuning language policy in Hong Kong education: stakeholders' perceptions, practices and challenges. *Language and Education*, 28(5), 459-476.
- Cheng, L. & C. DeLuca (2011). Voices from test-takers: Further evidence for test validation and test use. *Educational Assessment*, 16(2), 104–122.
- Cheng, L. (2002). The washback effect on classroom teaching of changes in public examinations. In S. J. Savignon (eds.), *Interpreting communicative language teaching: Contexts and concerns in teacher education*. New Haven, CT: Yale University Press, 91–111.
- Cheng, L. (2008). Washback, impact and consequences. In E. Shohamy & N. H. Hornberger (eds.), *Encyclopedia of Language and Education (2nd edn). Language Testing and Assessment* (vol. 7), New York, NY: Springer, 349–364.
- Cheng, L., S. Andrews & Y. Yu (2011). Impact and consequences of school-based assessment in Hong Kong: Views from students and their parents. *Language Testing*, 28(2), 221–250.
- Cheng, L., D. Klinger & Y. Zheng (2007). The challenges of the Ontario Secondary School Literacy Test for second language students. *Language Testing*, 24(2), 185–208.
- Cheng, L., Sun, Y., & Ma, J. (2015). Review of Washback Research Literature within Kane's Argument-Based Validation Framework. *Language Teaching*, 48(4), 436-470.

- Chik, A. & S. Besser (2011). International language test taking among young learners: A Hong Kong case study. *Language Assessment Quarterly*, 8(1), 73–91.
- Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.
- Creswell, J. W. (2015). *A concise introduction to mixed methods research*. Sage Publications.
- Curriculum Development Council. (2004). *English Language Curriculum Guide (Primary 1-6)*. Retrieved from http://www.edb.gov.hk/attachment/en/curriculum-development/kla/eng-edu/primary%201_6.pdf
- Curriculum Development Council. (2017). *English Language Education Key Learning Area Curriculum Guide (Primary 1-Secondary 6)*. Retrieved from http://www.edb.gov.hk/attachment/en/curriculum-development/kla/eng-edu/primary%201_6.pdf
- Curriculum Development Council., & Hong Kong Examinations and Assessment Authority. (2015). *English Language Curriculum and Assessment Guide (Secondary 4 - 6)*. Retrieved from http://www.edb.gov.hk/attachment/en/curriculum-development/kla/eng-edu/Curriculum%20Document/EngLangCAGuide_Nov15.pdf
- Dörnyei, Z. (2007). Research methods in applied linguistics.
- Evans, S. (2010). Business as usual: The use of English in the professional world in Hong Kong. *English for Specific Purposes*, 29(3), 153-167.
- Ferman, I. (2004). The washback of an EFL national oral matriculation test to teaching and learning. In L. Cheng, Y. Watanabe & A. Curtis (eds.), *Washback in language testing: Research contexts and methods*. Mahwah, NJ: Lawrence Erlbaum Associates, 191–210.
- Fields, A. (2013). *Discovering statistics using IBM SPSS statistics*. Thousand Oaks, CA.

- Forestier, K., Adamson, B., Han, C., & Morris, P. (2016). Referencing and borrowing from other systems: the Hong Kong education reforms. *Educational Research*, 58(2), 149-165.
- Frederiksen, J. R. and A. Collins. 1989. 'A systems approach to educational testing. *Educational Researcher*, 18(9), 27–32.
- Gibson, W. (2015). Qualitative Research Analysis in Education. In D. Hartas (ed.), *Educational research and inquiry: Qualitative and quantitative approaches*. Bloomsbury Publishing, 295-306.
- Gillham, B. (2008). *Developing a questionnaire*. A&C Black.
- Gorden, R. L. (1975). *Interviewing: Strategy, techniques, and tactics*. “The” Dorsey Press.
- Gosa, C. M. C. (2004). *Investigating washback: A case study using student diaries*. Ph.D. dissertation, Lancaster University.
- Green, A. (2006a). Washback to the learner: Learner and teacher perspectives on IELTS preparation course expectations and outcomes. *Assessing Writing*, 11, 113–134.
- Green, A. (2006b). Watching for washback: Observing the influence of the International English Language Testing System academic writing test in the classroom. *Language Assessment Quarterly*, 3, 333–368.
- Green, A. (2007a). *IELTS washback in context: Preparation for academic writing in higher education. Studies in Language Testing* 25. Cambridge: Cambridge University Press and Cambridge ESOL.
- Green, A. (2007b). Washback to learning outcomes: A comparative study of IELTS preparation and university pre-session language courses. *Assessment in Education*, 14, 75–97.
- Hartas, D. (Ed.). (2015). *Educational research and inquiry: Qualitative and quantitative approaches*. Bloomsbury Publishing.

- Hawkey, R. A. H. (2006). *Impact theory and practice: Studies of the IELTS test and ProgettoLingue 2000. Studies in Language Testing 24*. Cambridge: Cambridge University Press and Cambridge ESOL.
- Hayes, B. & J. Read (2004). IELTS test preparation in New Zealand: Preparing students for the IELTS academic module. In L. Cheng, Y. Watanabe & A. Curtis (eds.), *Washback in language testing: Research contexts and methods*. Mahwah, NJ: Lawrence Erlbaum Associates, 97–112.
- Hayes, B. M. (2003). *IELTS preparation in New Zealand: An investigation into the nature of the courses and evidence of washback*. Ph.D. dissertation, Victoria University of Wellington.
- Hong Kong Examinations and Assessment Authority. (2016). *2017 HKDSE English Language Assessment Framework*. Retrieved April 17, 2017, from http://www.hkeaa.edu.hk/DocLibrary/Media/Leaflets/HKDSE_SRR_A4_Booklet_Jun2011.pdf
- Hong Kong Examinations and Assessment Authority. (2017a). *2017 HKDSE Registration Statistics*. Retrieved April 17, 2017, from http://www.hkeaa.edu.hk/DocLibrary/Media/FactFigures/2017HKDSE_registrationstat.pdf
- Hong Kong Examinations and Assessment Authority. (2017b). *Recognition of Qualifications for HKDSE (Local)*. Retrieved April 17, 2017, from http://www.hkeaa.edu.hk/en/recognition/hkdse_recognition/local/
- Hong Kong Examinations and Assessment Authority. (2017c). *Recognition of Qualifications for HKDSE (Overseas)*. Retrieved April 17, 2017, from http://www.hkeaa.edu.hk/en/recognition/hkdse_recognition/ircountry_hkdse.html
- Hu, G. (2005). Contextual influences on instructional practices: A Chinese case for an ecological approach to ELT. *TESOL Quarterly*, 39(4), 635-660.

- Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Field methods*, 18(1), 3-20.
- Jolliffe, I. (2002). *Principal component analysis*. John Wiley & Sons, Ltd.
- Kan, V., & Adamson, B. (2010). Language policies for Hong Kong schools since 1997. *London Review of Education*, 8(2), 167-176.
- Kane, M. T. (2006). *Validation*. In R. L. Brennan (eds.), *Educational measurement (4th edn)*. Westport, CT: American Council on Education/Praeger, 17–64.
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement* 50.1, 1–73.
- Keith, T. Z. (2014). *Multiple regression and beyond: An introduction to multiple regression and structural equation modeling*. Routledge.
- King, N., & Horrocks, C. (2010). *Interviews in qualitative research*. Sage.
- Kunnan, A. J. (2004). Test fairness. In M. Milanovic, C. Weir & S. Bolton (eds.), *Europe language testing in a global context: Selected papers from the ALTE conference in Barcelona*. Cambridge: Cambridge University Press, 27–48.
- Kunnan, A. J. (2010). Test fairness and Toulmin's argument structure. *Language Testing*, 27(2), 183-189.
- Kvale, S. (1996). *An Introduction tot Qualitative Research Interviewing*.
- Latimer, D. (2009). *Washback effects of the Cambridge Preliminary English Test at an Argentinean bilingual school*. Ph.D. dissertation, University of Georgia.
- Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching*, 40(3), 243-249

- Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative research: A guide to design and implementation*. John Wiley & Sons.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13, 243–256.
- Mickan, P. & J. Motteram (2008). *An ethnographic study of classroom instruction in an IELTS preparation program*. www.ielts.org/pdf/Volume8_Report1.pdf.
- Mickan, P., & Motteram, J. (2009). The preparation practices of IELTS candidates: Case studies. *International English Language Testing System (IELTS) Research Reports 2009: Volume 10*, 1.
- Morris, P., & Adamson, B. (2010). *Curriculum, schooling and society in Hong Kong* (Vol. 1). Hong Kong University Press.
- Morrow, K. 1986. 'The evaluation of tests of communicative performance' in Portal (ed.). Paris, S. G., T. A. Lawton, J. C. Turner, and J. L. Roth. 1991. A developmental perspective on standardized achievement testing.' *Educational Researcher* 20(5), 12-19.
- Patton, M. Q. (2002). *Qualitative evaluation and research methods*. SAGE Publications, inc.
- Peacock, M. (1998). Comparing learner and teacher views on classroom activities for EFL. *International Journal of Applied Linguistics*, 8 (2), 233–250.
- Poon, A. Y. (2013). Will the new fine-tuning medium-of-instruction policy alleviate the threats of dominance of English-medium instruction in Hong Kong?. *Current Issues in Language Planning*, 14(1), 34-51.
- Qi, L. (2004). Has a high-stakes test produced the intended changes. *Washback in language testing: Research contexts and methods*, 171-190.
- Qi, L. (2005). Stakeholders' conflicting aims undermine the washback function of a high-stakes test. *Language Testing*, 22, 142–173.

- Rao, C., K. McPherson, R. Chand & V. Khan (2003). Assessing the impact of IELTS preparation programs on the General Training reading and writing test modules. *IELTS Research Report* (vol. 4). www.ielts.org/pdf/Vol5Report5.pdf.
- Read, J., & Hayes, B. (2003). The impact of IELTS on preparation for academic study in New Zealand. *International English Language Testing System (IELTS) Research Reports 2003: Volume 4*, 153.
- Saville, N. & R. Hawkey (2004). *The IELTS impact study: Investigating washback on teaching materials*. In Cheng et al. (eds.), 73–96.
- Schreiber, J. B., & Asner-Self, K. (2010). *Educational research*. Wiley Global Education.
- Shih, C. (2007). A new washback model of students' learning. *The Canadian Modern Language Review*, 64(1), 135–162.
- Shih, C. M. (2010). The washback of the General English Proficiency Test on university policies: A Taiwan case study. *Language Assessment Quarterly*, 7, 234–254.
- Shohamy, E. (1998). Critical language testing and beyond. *Studies in educational evaluation*, 24(4), 331–345.
- Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. Essex: Longman.
- Shohamy, E. (2007). Language tests as language policy tools. *Assessment in Education*, 14(1), 117–130.
- Shohamy, E., S. Donitsa-Schmidt & I. Ferman (1996). Test impact revisited: Washback effect over time. *Language Testing* 13, 298–317.
- Silverman, D. (2013). *Doing qualitative research: A practical handbook*. SAGE Publications Limited.

- Smart ,C., Drave, N.,& Shiu, J. (2014). Implementing innovation: a graded approach to English language testing in Hong Kong. In Coniam, D.& Benson, P. (ed.), *English language education and assessment: recent developments in Hong Kong and Chinese mainland*. DOI 10.1007/978-981-287-071-1_16.
- Stoneman, B. W. H. (2006). *The impact of an exit English test on Hong Kong undergraduates: A study investigating the effects of test status on students' test preparation behaviors*. Unpublished doctoral dissertation, Hong Kong Polytechnic University.
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. American Psychological Association.
- Tsagari, D. (2007). *Investigating the washback effect of a high-stakes EFL exam in the Greek context: Participants' perceptions, material design and classroom applications*. Unpublished doctoral dissertation, Lancaster University.
- Turner, C. (2006). Professionalism and high-stakes tests: Teachers' perspectives when dealing with educational change introduced through provincial exams. *TESL Canada Journal* 23, 54–76.
- Virkiru, L. I. (2011). From assessment to learning: The teaching of English beyond examinations. *Educational Forum*, 75, 129–142.
- Wall, D. & T. Horák (2006). The impact of changes in the TOEFL examination on teaching and learning in Central and Eastern Europe – Phase 1: The baseline study. *TOEFL Monograph Series, MS-34*. Princeton, NJ: Educational Testing Service.
- Wall, D. & T. Horák (2008). *The impact of changes in the TOEFL examination on teaching and learning in Central and Eastern Europe – Phase 2: Coping with change. (Report No. RR-08–37)*.
www.ets.org/Media/Research/pdf/RR-08-37.pdf.

- Wang, J. (2011). *A study of role of the 'teacher factor' in washback*. Unpublished doctoral dissertation, University of McGill, Montreal.
- Watanabe, Y. (1992). Washback effects of college entrance examination on language learning strategies. *JACET Bulletin*, 23, 175–194
- Watanabe, Y. (1996). Does grammar translation come from the entrance examination? Preliminary findings from classroom-based research. *Language Testing* 13.3, 318–333.
- Watanabe, Y. (2004). Methodology in washback studies. *Washback in language testing: Research contexts and methods*, 19-36.
- Xie, Q. (2015). Do component weighting and testing method affect time management and approaches to test preparation? A study on the washback mechanism. *System*, 50, 56-68.
- Xie, Q. & S. Andrews (2012). Do test design and uses influence test preparation? Testing a model of washback with structural equation modeling. *Language Testing*, 30(1), 49–70.
- Zhan, Y., & Andrews, S. (2014). Washback effects from a high-stakes examination on out-of-class English learning: insights from possible self theories. *Assessment in Education: Principles, Policy & Practice*, 21(1), 71-89.

Appendix 1 – Research Proposal

Examining Washback on Learning from a Sociocultural Perspective: The Case of a Graded Approach to English Language Testing in Hong Kong

1. Introduction

Nearly two decades of research into washback in the field of applied linguistics beginning with Alderson and Wall's (1993) landmark seminal work has given rise to an expanding body of empirical washback literature. Over the years, parallel to the ever-increasing prevalence of large-scale, high-stakes, and often international, testing worldwide (Cheng, 2008; Green, 2006b; Wall & Horák, 2006, 2008) and consequently the growing awareness of the enormous power testing has on the stakeholders (Kunnan, 2004; Shohamy, 2001, 2007), the conceptualization of washback has shifted from, initially, an aspect of construct validity (Messick, 1996), to a free-standing construct founded upon the intricate relationships between testing, teaching, and learning (Bailey, 1996; Cheng, 2002; Green, 2007), and, increasingly, to a socially-situated construct which transcends the boundary between washback and impact by bringing together a range of stakeholders and factors within and beyond the classroom context (Booth, 2012; Shih, 2007, 2010). In other words, both the conception and scope of washback has evolved over time. Thanks to this, washback studies encompassing a diverse range of areas such as validation of test use (e.g. Bachman, 2005; Kane, 2006, 2013), impact evaluation (e.g. Saville & Hawkey, 2004; Green, 2006b; Wall & Horák, 2006, 2008), and critical language testing (e.g. Kunnan, 2004; Shohamy, 2001, 2007), and across a wide range of contexts including Canada (e.g. Cheng, Klinger & Zheng, 2007), China (e.g. Xie & Andrews, 2012), and Sri Lanka (Wall & Alderson, 1993) have emerged and flourished in the field.

Among the many areas investigated in empirical work on washback stand the impact studies which examine the systematic changes to teaching and learning brought about by testing (Cheng, Sun & Ma, 2015). While earlier impact studies focus largely on washback on teaching and have offered ample evidence of how testing influences aspects of teaching including teachers' attitude (e.g. Wang, 2011), content of teaching (e.g. Hayes & Read, 2004; Qi, 2005), and teaching methods (e.g. Mickan & Motteram, 2008), in recent years washback on learning has also begun to receive attention. For example, studies tapping into the washback effects among learners (e.g. Green, 2006a; Hawkey, 2006) and the mediating factors shaping their washback (e.g. Cheng & Deluca, 2011; Xie & Andrews, 2012) have been conducted to explicate this complicated process. Yet, although these studies have advanced the understanding of washback, the majority of them have either focused rather narrowly on the observable washback effects (e.g. Green, 2006a; Latimer, 2009) and/or on the mediating factors at the individual level in isolation of the wider sociocultural context (e.g. Cheng, Andrews & Yu, 2011; Cheng & Deluca, 2011). In other words, there is, I argue, insufficient attention paid to the fact that washback on learning is not an individual but a socially-situated phenomenon (Booth, 2012; Shih, 2007) under which learners' observable actions are driven by an array of mediating factors not only within but also beyond the individual level. There is, therefore, a need to go further and investigate washback on learning more thoroughly and holistically.

In response to this need, this proposed study is a modest attempt to approach washback on learning from a sociocultural perspective. Drawing upon multiple interviews and a questionnaire, this sequential mixed methods study seeks to examine the washback effects that Hong Kong learners perceive

to have following the introduction of the new graded approach in the Hong Kong Diploma of Secondary Education English Language Examination (HKDSE-English). In particular, this study attempts to uncover students' self-perceived washback effects, and the full range of individual and sociocultural mediating factors shaping such effects. The graded approach in HKDSE-English implemented in 2012, which in short allows test-takers to choose between an easier section and a more difficult section in the test, is chosen as the subject of interest because of two primary reasons. First, the washback of a graded approach to language testing has yet to be empirically investigated (Smart, Drave & Shiu, 2014) and is thus deemed to be examined. Second, and more importantly, given the examination-oriented culture (Berry, 2011), the stress on meritocracy (Brown et al, 2009), and the fact that the HKDSE English Language Examination is a high-stakes test; the test is likely to impose significant washback effects on Hong Kong learners which in turn offers a rich case of analysis. Through carefully investigating the issue from a sociocultural perspective, this study aims at enriching the understanding of washback on learning.

2. Literature Review

As highlighted in the above discussion, it is not until recent years when washback on learning, unlike washback on teaching which has been comparatively more well-researched (Cheng, Sun & Ma, 2015), has been empirically investigated in impact studies. In this literature review, the developing body of studies on the topic is presented and discussed in order to underscore, first, the fact that washback on learning is still a developing area, and, subsequently, the gaps that are to be filled if it is to be understood more thoroughly and holistically. Mirroring the two major lines of research on the topic, this review is structured in two major sub-sections: 1) studies on the observable washback effects on learning, and 2) studies on the underlying mediating factors which shape learners' washback.

2.1 Studies on the Observable Washback Effects on Learning

The earlier phase of impact studies investigating washback on learning (e.g. Hawkey, 2006; Hayes & Read, 2004; Gosa, 2004; Green, 2006a) focuses primarily on the direct, observable washback effects high-stakes testing has on learning which are, in most cases, realized in the form of learners' preference and participation in extensive test preparation activities. Hawkey (2006)'s classroom observation study which examines the impact of IELTS is a classic example. As concluded by Hawkey (2006), learners' demand for a test-preparation course instead of a general English course, as realized by how they twist the focus of the course from a balanced curriculum to a test-focused curriculum against the teachers' preference, is a clear indicator of washback on learning. This finding, which is supported by several studies (Gosa, 2004; Hayes, 2003; Latimer, 2009) which all recognize the power learners have in shaping language classes, asserts that a narrowed, restricted curriculum resulted from learners' preference for test-preparation activities in class is one of the direct washback effects on learning.

Apart from the classroom context, previous studies (Gosa, 2004; Mickan & Motteram, 2009; Tzagari, 2009; Stoneman, 2006) have also indicated that such test-preparation activities extend beyond classrooms to learners' personal learning environment. As suggested by Gosa (2004)'s diary study which examines the washback effects the Romanian school-leaving examination has on a group of secondary students, learners often participate in intensive test-preparation activities including coaching and cramming outside the schools whenever their demand for test-specific drilling is not fulfilled in classes.

Gosa's (2004) finding, echoing that of the studies conducted in Australia (Mickan & Motteram, 2009), Greece (Tsayari, 2009), and Hong Kong (Stoneman, 2006), further elaborates on the range of actions learners might take in response to high-stakes testing.

In addition to digging into the washback effects on learners' actions, a small body of work (e.g. Green, 2007; Latimer, 2009; Virkuru) turns to the effects on learners' learning outcome. While studies into the effectiveness of test-preparation courses have yet to reach consensus (e.g. Rao, McPherson, Chan & Kwan, 2003; Green, 2007), what appears to be agreed upon is the fact that washback on learning leads to both positive and negative impact on learners' learning outcome, a view in line with Bailey (1996) and Alderson & Wall's (1993) hypotheses. On the one hand, testing brings about improved learning outcome in that tested skills are taught, studied, and revised as realized in the score gains (Rao et al, 2003). On the other hand, testing simultaneously discourage the study of untested skills (Virkuru, 2011), and thereby lead to a narrowed set of learning outcome specific to the test (Latimer, 2009).

In short, what this line of earlier research has taught us is that washback on learning is a complex and multi-faceted construct that is fundamentally different from, and might even exert impact on, washback on teaching (e.g. Hawkey, 2006), and is, therefore, an area of research in its own right. Yet, while acknowledging their contributions, what should be noted is that the majority of these studies, often focusing only on describing the observable aspects of washback (i.e. learners' actions and learning outcomes) generalized from their limited data set, are rather insensitive to the reasons/factors underneath the observable outcome and, more importantly, the fact that washback on learning is not homogeneous among the learners (Cheng, Sun & Ma, 2015). It is, indeed, a growing awareness towards this inadequacy that drives and powers the second phase of studies into washback on learning.

2.2 Studies on the Underlying Mediating factors that Shape Learners' Washback

Realizing the inadequacy of not going beyond the observable aspects, the subsequent phase of research, guided by the assumption that washback is different across individual learners, has shifted from a solely descriptive approach to an exploratory approach which scrutinizes the sea of meditating factors that shape learners' washback. In general terms, these factors fall under two broad categories: intrinsic factors (i.e. at the individual level) and extrinsic factors (i.e. beyond the individual level).

One of the first intrinsic mediating factors identified in the literature is learners' language proficiency (Cheng, Andrews & Yu, 2011; Ferman, 2004; Xie and Andrews, 2012). In Ferman's (2004) study, Israeli learners with lower proficiency are found to participate in a series of test-preparation activities including intensive drilling and individual coaching more actively than their more competent peers in preparation for the national speaking examination. This finding is corroborated by several subsequent studies (e.g. Cheng, Andrews & Yu, 2011; Xie & Andrews, 2012) which all draw correlations between learners' language proficiency and their intensity of test preparation. In addition to learners' language proficiency, what Xie and Andrews's (2012) large-scale quantitative study conducted in China uncover are two other significant intrinsic mediating factors: learners' perceptions of test design and their perceptions of test use. Xie and Andrews's (2012) study, therefore, have added personal expectancy and motivation, both supported by later studies (e.g. Booth, 2012; Zhan & Andrews, 2014), to the scene where an increasing number of intrinsic factors such as learners' previous experience (Cheng & DeLuca, 2011; Zhan & Andrews, 2014) has gradually come into play.

Meanwhile, a few researchers (e.g. Chik & Besser, 2011; Zhan & Andrews, 2014) are increasingly aware of the existence of extrinsic mediating factors which are equally powerful in shaping learners' washback. Zhan and Andrews (2014), for example, identify 'learning environment' and 'others' experiences in taking the test', under which a full range of stakeholders and objects are included, as common mediating factors influencing "what they (learners) learn then how they learn" (p71). Similarly, Chik and Besser (2011) also outline the influences of schools, private learning centres, and the media in their discussion of how Hong Kong young learners' learning are shaped by the interaction of extrinsic factors under the prevalence of international testing. The two studies, thus, represents the growing recognition of the power of extrinsic factors in mediating learners' learning.

Interestingly, instances when intrinsic and extrinsic factors are studied together are scarce. Shih (2007)'s study in Taiwan which reveals how washback on learners' learning is mediated by the interaction between an array of "intrinsic factors, extrinsic factors, and test factors" (p151) is one of the very few examples. Shih's (2007) study together with that of Booth (2012), although on a small scale, offer a much-needed perspective into how washback can be approached from a socially-situated view.

To sum up the current phase of research into washback on learning, the body of growing research has gone beyond the mere description of the observable aspects to a close examination of the intrinsic and extrinsic mediating factors influencing learners' washback. Yet, this is not to say the work is comprehensive enough to account for the complexity of washback. In particular, given how studies on intrinsic factors outnumber that of extrinsic factors, mediating factors beyond the individual level have yet to receive sufficient attention. Furthermore, the fact that most factors, especially the intrinsic ones, are often studied independently in isolation of the wider socio-cultural context, and that research into intrinsic factors and extrinsic factors appear to go parallel to one another, seem to suggest that washback has still not been considered as a socially-situated phenomenon, a view that is increasingly advocated in the wider field of applied linguistics (Cheng, Sun & Ma, 2015).

2.3 The Need of Studying Washback on Learning from a Sociocultural Perspective

It should now become clear that there is a need to approach washback on learning from a sociocultural perspective which bridges together the discrete elements in the field: the observable washback effects, intrinsic mediating factors, and extrinsic mediating factors. The present study, which examines collectively the washback effects that Hong Kong secondary school learners perceive to have following the introduction of the graded approach in HKDSE-English, and the range of underlying intrinsic and extrinsic mediating factors shaping such self-perceived washback effects, is a modest attempt to address the need. The main research questions of the study are as follows:

1. What are the washback effects that Hong Kong secondary school learners perceive to have on their learning following the introduction of the graded approach in HKDSE-English?
2. What are the possible intrinsic and extrinsic mediating factors that shape learners' self-perceived washback effects?

3. Methods

3.1 Theoretical Framework

The theoretical framework that guides the present study is an adapted version of Shih's (2007) "Washback Model of Student's Learning" (p151). In this model (see Appendix 1), washback on learning is conceptualized as a socially-situated phenomenon in which the outcome, "washback of a test on students' learning and psychology" (p151) is dynamically negotiated by an array of "extrinsic factors, intrinsic factors, and test factors" (p151). As will be discussed in greater detail in the following sections, this model is adopted as the template for the design of research instruments in this study.

3.1 Targeted Participants

The targeted participants are secondary four to six Hong Kong ESL learners who are currently studying in local secondary schools and are expected to undertake HKDSE-English within two years. 12 participants are expected in phase one of the study while 120 participants are expected in phase two. For both phases, maximum-variation sampling which captures "the widely varying instances of the phenomenon" (Merriam & Tisdell, 2015, p79), will be used to recruit participants from all three bandings of the schools in Hong Kong which are categorized according to students' academic ability.

3.2 Instruments and Procedure

The proposed sequential mixed-methods study comprises two main phases. Phase one, which aims at developing a questionnaire that includes a wide range of self-perceived washback effects and mediating factors that exist among Hong Kong learners, is qualitative in nature. Using Shih's (2007) broad categories of possible washback effects and mediating factors (see Appendix 1) as an template for the development of interview questions, three 20-minute semi-structured focus-group interviews with the targeted participants (N=4 in each group) are conducted to elicit as many possible self-perceived washback effects and mediating factors as possible. Following the interviews, such effects and mediating factors identified are then rephrased into statements which are to be put on a 5-point Likert-scale questionnaire. From then onwards, the questionnaire, which is to be administered to a wider population in the subsequent phase, enters a validation process in which the questionnaire is pilot-tested with 15 learners. Amendments regarding issues such as the skewness of the data and item distribution will be made in order to ensure the "clarity, accuracy, and appropriacy" (Xie, 2015, p60) of the items.

Phase two commences with the administration of the questionnaire to 120 learners. The aim of this is to screen for the self-perceived washback effects and mediating factors that are under-represented in the wider population so that 1) the prevalent self-perceived washback effects among Hong Kong learners, and 2) the major groups of mediating factors shaping learners' self-perceived washback effects, can be identified to answer the RQs.

3.2 Data Analysis

The present study involves both qualitative (focus-group-interviews) and quantitative data (questionnaire). The interviews are to be analyzed using an inductive approach in which themes emerge from the data (Silverman, 2013). As for the questionnaire, it is to be analyzed using both descriptive statistics and principal component analysis which identifies the clustering of factors (Jolliffe, 2002).

3.3. *Ethics*

As the present study involves participants below the age of 18, both written consent from schools and parents, and verbal consent from the students are needed before the collection of data. Furthermore, during the course of study, all the participants are to be kept strictly anonymous and they reserve the right to withdraw from the study at any time. All data will be stored securely in locked computers and drives at all times.

4. Conclusion

To sum up, the present study examines collectively students' self-perceived washback effects following the introduction of the graded approach in HKDSE-English and the range of underlying intrinsic and extrinsic mediating factors shaping such effects. Adopting a sociocultural approach which is much needed in the field, this study seeks to contribute to the field by deepening the understanding of both the washback effects of a graded approach to language testing, and, on a more macroscopic level, washback on learning by bridging the observable washback effects, intrinsic mediating factors, and extrinsic mediating factors.

5. Bibliography

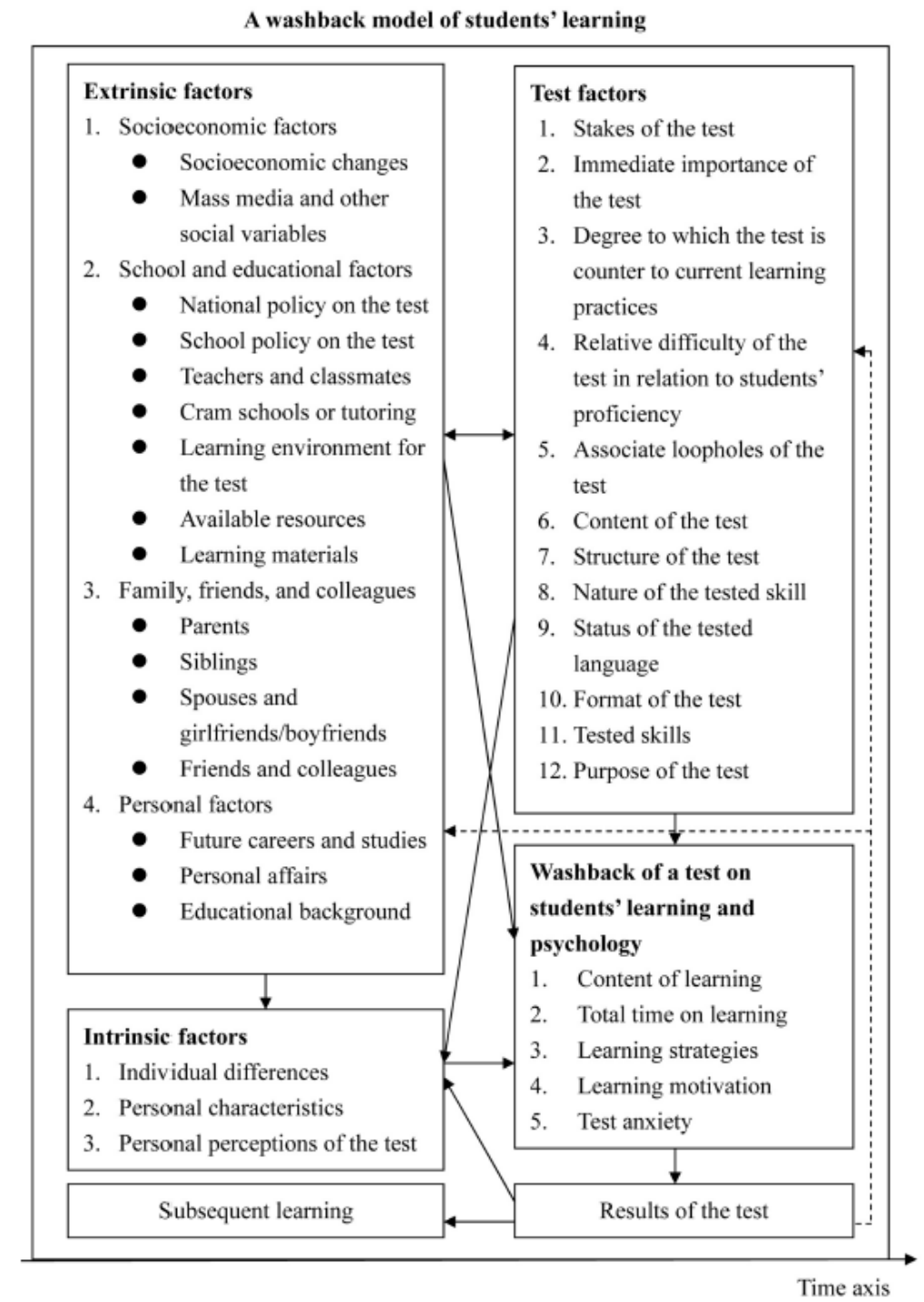
- Alderson, J. C. & D. Wall (1993). 'Does washback exist?' *Applied Linguistics*, 14, 115–129.
- Bachman, L. F. (2005). Building and supporting a case for test use. *Language Assessment Quarterly*, 2(1), 1–34.
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13, 257–279.
- Berry, R. (2011). Assessment Trends in Hong Kong: Seeking to Establish Formative Assessment in an Examination Culture. *Assessment in Education: Principles, Policy & Practice*, 18(2), 199–211.
- Booth, D. K. (2012). *Exploring the washback of the TOEIC in South Korea: a sociocultural perspective on student test activity*. Ph.D. dissertation, The University of Auckland.
- Brown, G. L., Kennedy, K. J., Fok, P. K., Chan, J. S., & Yu, W. M. (2009). Assessment for Student Improvement: Understanding Hong Kong Teachers' Conceptions and Practices of Assessment. *Assessment in Education: Principles, Policy & Practice*, 16(3), 347–363.
- Cheng, L. & C. DeLuca (2011). Voices from test-takers: Further evidence for test validation and test use. *Educational Assessment*, 16(2), 104–122.
- Cheng, L. (2002). The washback effect on classroom teaching of changes in public examinations. In S. J. Savignon (eds.), *Interpreting communicative language teaching: Contexts and concerns in teacher education*. New Haven, CT: Yale University Press, 91–111.
- Cheng, L. (2008). Washback, impact and consequences. In E. Shohamy & N. H. Hornberger (eds.), *Encyclopedia of Language and Education (2nd edn)*. *Language Testing and Assessment* (vol. 7), New York, NY: Springer, 349–364.
- Cheng, L., S. Andrews & Y. Yu (2011). Impact and consequences of school-based assessment in Hong Kong: Views from students and their parents. *Language Testing*, 28(2), 221–250.
- Cheng, L., D. Klinger & Y. Zheng (2007). The challenges of the Ontario Secondary School Literacy Test for second language students. *Language Testing*, 24(2), 185–208.
- Cheng, L., Sun, Y., & Ma, J. (2015). Review of Washback Research Literature within Kane's Argument-Based Validation Framework. *Language Teaching*, 48(4), 436–470.
- Chik, A. & S. Besser (2011). International language test taking among young learners: A Hong Kong case study. *Language Assessment Quarterly*, 8(1), 73–91.
- Ferman, I. (2004). The washback of an EFL national oral matriculation test to teaching and learning. In L. Cheng, Y. Watanabe & A. Curtis (eds.), *Washback in language testing: Research contexts and methods*. Mahwah, NJ: Lawrence Erlbaum Associates, 191–210.
- Gosa, C. M. C. (2004). *Investigating washback: A case study using student diaries*. Ph.D. dissertation, Lancaster University.

- Green, A. (2006a). Washback to the learner: Learner and teacher perspectives on IELTS preparation course expectations and outcomes. *Assessing Writing*, 11, 113–134.
- Green, A. (2006b). Watching for washback: Observing the influence of the International English Language Testing System academic writing test in the classroom. *Language Assessment Quarterly*, 3, 333–368.
- Green, A. (2007a). *IELTS washback in context: Preparation for academic writing in higher education. Studies in Language Testing* 25. Cambridge: Cambridge University Press and Cambridge ESOL.
- Green, A. (2007b). Washback to learning outcomes: A comparative study of IELTS preparation and university pre-session language courses. *Assessment in Education*, 14, 75–97.
- Hawkey, R. A. H. (2006). *Impact theory and practice: Studies of the IELTS test and ProgettoLingue 2000. Studies in Language Testing* 24. Cambridge: Cambridge University Press and Cambridge ESOL.
- Hayes, B. & J. Read (2004). IELTS test preparation in New Zealand: Preparing students for the IELTS academic module. In L. Cheng, Y. Watanabe & A. Curtis (eds.), *Washback in language testing: Research contexts and methods*. Mahwah, NJ: Lawrence Erlbaum Associates, 97–112.
- Hayes, B. M. (2003). *IELTS preparation in New Zealand: An investigation into the nature of the courses and evidence of washback*. Ph.D. dissertation, Victoria University of Wellington.
- Jolliffe, I. (2002). *Principal component analysis*. John Wiley & Sons, Ltd.
- Kane, M. T. (2006). *Validation*. In R. L. Brennan (eds.), *Educational measurement (4th edn)*. Westport, CT: American Council on Education/Praeger, 17–64.
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement* 50.1, 1–73.
- Kunnan, A. J. (2004). Test fairness. In M. Milanovic, C. Weir & S. Bolton (eds.), *Europe language testing in a global context: Selected papers from the ALTE conference in Barcelona*. Cambridge: Cambridge University Press, 27–48.
- Latimer, D. (2009). *Washback effects of the Cambridge Preliminary English Test at an Argentinean bilingual school*. Ph.D. dissertation, University of Georgia.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13, 243–256.
- Mickan, P. & J. Motteram (2008). *An ethnographic study of classroom instruction in an IELTS preparation program*. www.ielts.org/pdf/Volume8_Report1.pdf.
- Qi, L. (2005). Stakeholders' conflicting aims undermine the washback function of a high-stakes test. *Language Testing*, 22, 142–173.
- Rao, C., K. McPherson, R. Chand & V. Khan (2003). Assessing the impact of IELTS preparation programs on the General Training reading and writing test modules. *IELTS Research Report (vol. 4)*. www.ielts.org/pdf/Vol5Report5.pdf.

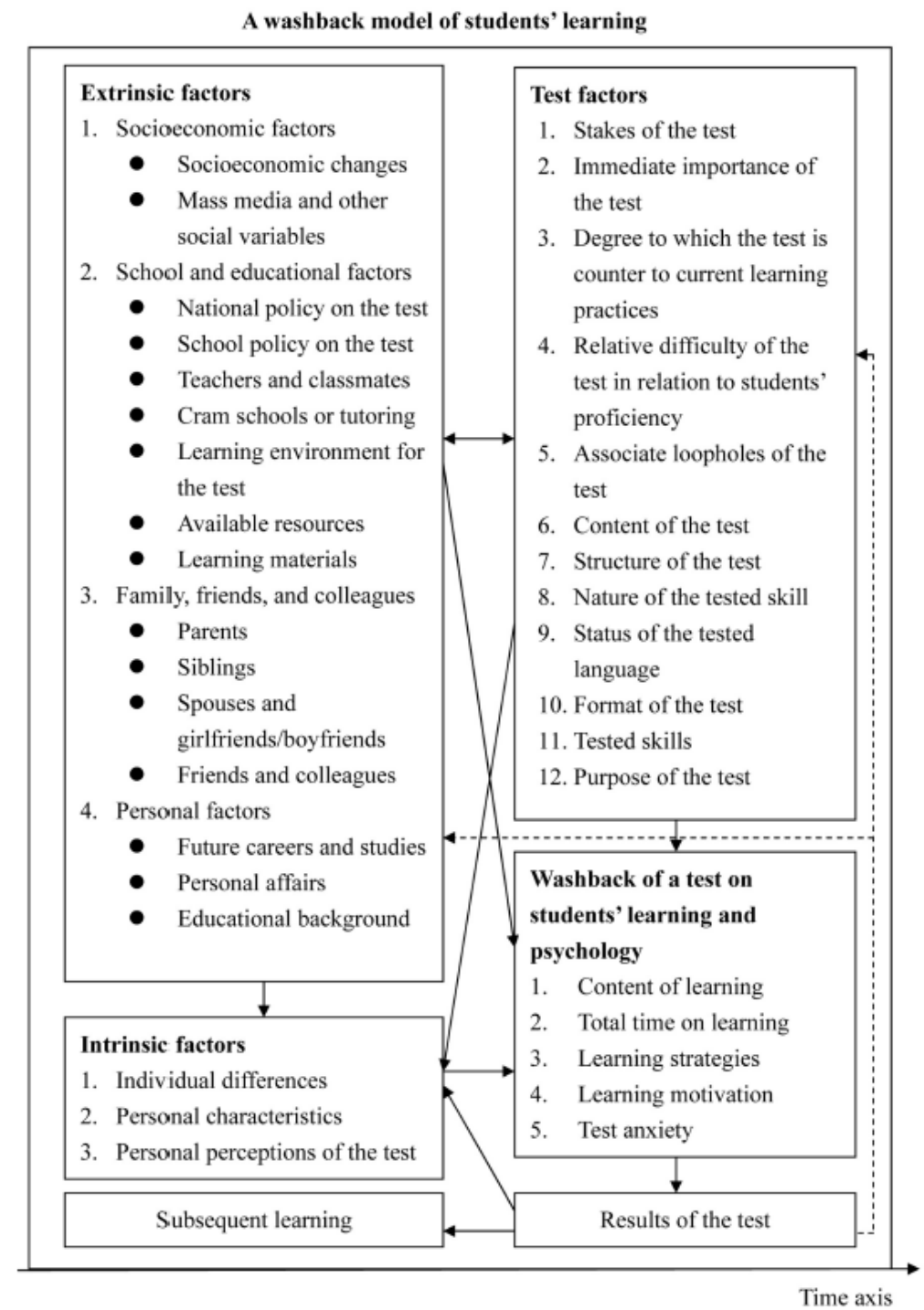
- Saville, N. & R. Hawkey (2004). *The IELTS impact study: Investigating washback on teaching materials*. In Cheng et al. (eds.), 73–96.
- Shih, C. (2007). A new washback model of students' learning. *The Canadian Modern Language Review*, 64(1), 135–162.
- Shih, C. M. (2010). The washback of the General English Proficiency Test on university policies: A Taiwan case study. *Language Assessment Quarterly*, 7, 234–254.
- Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. Essex: Longman.
- Shohamy, E. (2007). Language tests as language policy tools. *Assessment in Education*, 14(1), 117–130.
- Smart, C., Drave, N., & Shiu, J. (2014). Implementing innovation: a graded approach to English language testing in Hong Kong. In Coniam, D. & Benson, P. (ed.), *English language education and assessment: recent developments in Hong Kong and Chinese mainland*. DOI 10.1007/978-981-287-071-1_16.
- Stoneman, B. W. H. (2006). *The impact of an exit English test on Hong Kong undergraduates: A study investigating the effects of test status on students' test preparation behaviors*. Unpublished doctoral dissertation, Hong Kong Polytechnic University.
- Tsagari, D. (2007). *Investigating the washback effect of a high-stakes EFL exam in the Greek context: Participants' perceptions, material design and classroom applications*. Unpublished doctoral dissertation, Lancaster University.
- Virkiru, L. I. (2011). From assessment to learning: The teaching of English beyond examinations. *Educational Forum*, 75, 129–142.
- Wall, D. & T. Horák (2006). The impact of changes in the TOEFL examination on teaching and learning in Central and Eastern Europe – Phase 1: The baseline study. *TOEFL Monograph Series, MS-34*. Princeton, NJ: Educational Testing Service.
- Wall, D. & T. Horák (2008). *The impact of changes in the TOEFL examination on teaching and learning in Central and Eastern Europe – Phase 2: Coping with change. (Report No. RR-08–37)*. www.ets.org/Media/Research/pdf/RR-08-37.pdf.
- Wang, J. (2011). *A study of role of the 'teacher factor' in washback*. Unpublished doctoral dissertation, University of McGill, Montreal.
- Xie, Q. (2015). Do component weighting and testing method affect time management and approaches to test preparation? A study on the washback mechanism. *System*, 50, 56–68.
- Xie, Q. & S. Andrews (2012). Do test design and uses influence test preparation? Testing a model of washback with structural equation modeling. *Language Testing*, 30(1), 49–70.

Zhan, Y., & Andrews, S. (2014). Washback effects from a high-stakes examination on out-of-class English learning: insights from possible self theories. *Assessment in Education: Principles, Policy & Practice*, 21(1), 71-89.

Appendix 1 – A Washback Model of Student's Learning (Shih, 2007, p151)



Appendix 2 – A Washback Model of Student’s Learning (Shih, 2007, p151)



Appendix 3 – Interview Guide

The Opening of the Interview 歡迎詞	Purpose
<p>Good morning. Welcome to this interview. My name is Heskey and I am a student researcher at the Institute of Education, University College London. 早晨，歡迎參加這次訪問。我叫 Heskey 曾智豐。我是英國 UCL 倫敦大學學院教育學院的學生研究員。</p>	<p>To introduce the researcher</p>
<p>I am now working on a project about the influences that the new graded approach in HKDSE-English has on Hong Kong secondary school students. I hope to know through this project what the influences are and what causes these influences. This is exactly why I have invited you to this interview. Thank you very much for coming and your contribution is really very much appreciated. 我現在正在進行一項關於香港中學文憑試，即 HKDSE，英文科中新採用的 graded approach 對香港中學生的影響的研究。我希望通過這個研究知道究竟它產生了什麼影響和是什麼導致了這些影響，即是其起因。這正是我今次邀請大家來參與這次訪問的目的。非常感謝你們的到來和幫忙。</p>	<p>To contextualize the interview and to define the situation for the participants</p>
<p>There are three main parts in this interview. In part 1, I am interested in the feelings you have with the graded approach – what you think about Part B1 and Part B2. In part 2, I wish to find out what actions you take to prepare yourself for the graded approach. In part 3, I hope to find out what the motivations behind these actions are. For each part, I have outlined a list of questions. We will go through them and please share with me what you think and, more importantly, discuss these with your one another. Discussions are most welcome here. 這個訪問有三部分。第一部分，我想了解一下你對這個 graded approach 的感覺和看法。第二部分，我希望知道為了應對 HKDSE 中的 graded approach 你做了什麼具體的準備工作。第三部分，我希望了解你們這些行動背後的動力，是什麼激勵或者推進你們做這些準備工作。我已經針對每一個部分列出一系列的問題。我們會逐一探究，請大家都積極去分享自己的看法。更重要的是，大家一起討論一下。</p>	<p>To introduce what will happen and what is to be expected in the interview</p>
<p>Before we start, let me assure you that your identity will be kept strictly anonymous. This means that everything you say in this interview will not be attributed to you because I will use fake names such as Student A and Student B afterwards. Furthermore, please allow me to remind that you have the right to withdraw from this interview and the project at any time. You can do so by remaining silent, leaving the room, or speaking to me after the interview. If you choose to do so, I promise that all your data will be erased. 在開始之前，我會說明一下，在這個研究中，你們的身份會被保密，匿名的。即是你們在這個訪問中所說的內容都不會指名道姓說是你們說的，我會採用代稱，學生 A 學生 B 等。另外，你有權在這個研究和訪問中途任何時候提出退出參與。你只需要保持安靜，離開這間房，之後再告訴我即可。如果你決定退出，我保證你的所有資料和分享的內容都會被刪除。</p>	<p>To make explicit that full anonymity and the right to withdrawal are guaranteed</p>
<p>Is there anything that you would like to ask? (Pause) Now, may I check for one more time if you agree to participate in this interview?</p>	<p>To obtain verbal consent</p>

在訪問開始之前，你們還有什麼疑問需要我現在解釋一下？（停頓）現在，我想再次確認一下，是否每一位都同意參加這次訪問？	
---	--

Question	Question Type	Purpose	RQ
<i>Part 1 – Students' Perceptions towards the Graded approach in HKDSE-English</i> 第一部分-學生對HKDSE 英文科中 graded approach 的觀感			
1. Since 2012, the graded approach has been introduced in HKDSE-English. This means that section B of the reading paper and the listening & integrated skills paper is now divided into Part B1, the easier section, and Part B2, the more difficult section. How do you feel about Part B1 and Part B2? 自 2012 年起，香港 DSE 英文科中引入 graded approach，即是閱讀卷和聆聽及綜合運用卷的乙部(section B)中分為較為容易的 B1，和較難的 B2。你對考試分 B1、B2 有什麼看法？ Probes a) Which candidate would be most likely to choose Part B1? 你認為哪類考生會很大機會選考 B1？ b) Which candidate would be most likely to choose Part B2? 你認為哪類考生則很大機會選考 B2？ c) What are the characteristics of the people who choose B1? 你認為選考 B1 的考生有什麼特點或者特質？ d) What are the characteristics of the people who choose B2? 你認為選考 B2 的考生有什麼特點或者特質？ e) Why would somebody choose Part B1? 你認為為什麼某些人會選擇 B1？ f) Why would somebody choose Part B2? 你認為為什麼另一部分人會選擇 B2？	Exploratory, introductory, feeling, direct Follow-up, specifying, feeling, direct	To elicit learners' perception towards the graded approach To prepare the participants for the interview	2
<i>Part 2 – Learners' Perceived Washback Effects following the Introduction of the Graded Approach</i> 第二部分-學生認為的 graded approach 帶來的影響			
2. So we have largely agreed that Part B1 is (a summary of what Part B1 entails based on the responses) and Part B2 is (a summary of what Part B2 entails based on the responses). Now, let us move on to how you prepare for the graded approach. What do you do to prepare yourself for the reading paper and the listening & integrated skills paper in HKDSE-English? 好，我們現在大致同意 B1 是...(根據受訪者的回答總結 B1)和 B2 是...(根據受訪者的回答總結 B2)。現在，我們去下一部分就是關於你們如何因應 graded approach 去做準備，即備考。你為了準備 DSE 英文科閱讀卷和聆聽及綜合運用卷具體做了什麼？ Probes	Exploratory, introductory, experience/behaviour, direct	To elicit a list of learners' perceived washback effects	1

<p>a) What is the work you do in class (e.g. lessons; class practice) to prepare yourself for Part B1/ B2 of the reading paper? 你在課堂中做了什麼工作來準備閱讀卷的 B1 或 B2 部分呢?</p> <p>b) What is the work you do out-of-class (e.g. tutorial classes; personal time) to prepare yourself for Part B1/B2 of the reading paper? 你在課堂外，例如在補習社和課餘私人時間，又做了什麼工作來準備閱讀卷的 B1 或 B2 部分呢?</p> <p>c) What skills or areas do you practice most when you are preparing for the reading paper? 當你準備閱讀卷的時候，哪些具體的技能或者內容你會去練習多呢?</p> <p>d) What is the work you do in-class to prepare yourself for Part B1/B2 of the listening & integrated skills paper? 你在課堂內做了什麼工作來準備聆聽及綜合運用卷的 B1 或 B2 部分呢?</p> <p>e) What is the work you do out-of-class to prepare yourself for Part B1/B2 of the listening & integrated skills paper? 你在課堂外又做了什麼工作來準備聆聽及綜合運用卷的 B1 或 B2 部分呢?</p> <p>f) What skills or areas do you practice most when you are preparing for the listening & integrated skills paper? 當你準備聆聽及綜合運用卷的時候，哪些具體的技能或者內容你會去練習多呢?</p>	Follow-up, specifying, experience/behaviour, direct		
<p><i>Part 3 – Learners’ Self-reported Mediating Factors</i> 第三部分-學生自我反映的影響因素</p>			
<p>3. So the preparatory work you do for Part B1/B2 includes (a summary of the preparatory work). Now, I would like to know what motivates you to do so. On a personal level, what motivates you to do the preparatory work for Part B1/B2? 所以你們為了準備 B1/B2 部分你們做了...(總結前述的準備工作)。現在，我想了解一下，是什麼動機或者原因促使你們這樣做呢？個人層面上，是什麼促使你們去為 B1/B2 做這些準備工作？</p> <p>Probes</p> <p>a. What do you think of the role personal qualities or characteristics play? 你認為個人的性格特質有無影響或者促使你做這些準備工作呢？</p> <p>b. What do you think of the role personal ambitions play? 你認為個人目標又是否對此產生影響呢？</p> <p>c. What do you think of the role your language ability plays? 你認為你的語言能力又是否帶來影響呢？</p>	<p>Exploratory, introductory, opinion/value, direct</p> <p>Follow-up, specifying, opinion/value, direct</p>	To elicit a range of intrinsic mediating factors perceived by the learners	2

<p>4. Apart from the factors at the personal level such as (a summary of personal factors), are there any factors at the school level which influence your preparatory work for Part B1/B2? 除了個人層面的因素...(總結前述的個人因素), 還有什麼學校層面的因素影響你對 B1/B2 的準備工作呢?</p> <p>Probes</p> <p>a. What do you think of the role your friends play? 你認為你的朋輩是否對你的準備工作帶來影響?</p> <p>b. What do you think of the role your teachers play? 你認為你的老師又是否有對你的準備工作有影響?</p> <p>c. What do you think of the role other members at the school play? 你認為學校裡面的其他人員又是否對你有影響?</p>	<p>Exploratory, introductory, opinion/value, direct</p> <p>Follow-up, specifying, opinion/value, direct</p>	<p>To elicit a range of extrinsic mediating factors perceived by the learners</p>	<p>2</p>
<p>5. Right. So we have covered the factors from the school including (a summary of the school-level factors). Now, let's move on to the family. What would be the possible influences from the family which influence your preparatory work for Part B1/B2? 好, 所以我們已經討論了學校層面上的影響因素, 包括...(總結前述的學校層面因素)。現在, 我們看看家庭層面, 在家庭層面上有什麼可能的因素會影響你對 B1/B2 的準備呢?</p> <p>Probes</p> <p>a. What do you think of the role your parents play? 你認為你的父母對你的準備工作有無影響?</p> <p>b. What do you think of the role your siblings play? 你認為你的兄弟姐妹又是否帶來影響?</p> <p>c. What do you think of the role your relatives play? 你認為親戚又是否對你的準備工作帶來影響?</p>	<p>Exploratory, introductory, opinion/value, direct</p> <p>Follow-up, specifying, opinion/value, direct</p>	<p>To elicit a range of extrinsic mediating factors perceived by the learners</p>	<p>2</p>
<p>6. So the influencing factors from the family are (a summary of the factors identified). Last but not least, are there any factors at the societal level which influence your preparatory work for Part B1/B2? 所以你認為家庭層面的影響因素是有...(總結前述的因素)。最後, 我們看看社會層面, 社會上又是否有一些因素影響到你對 B1/B2 的準備呢?</p> <p>Probes</p> <p>a) What do you think of the role media plays? 你認為媒體是否對你的準備工作帶來影響?</p>	<p>Exploratory, introductory, opinion/value, direct</p> <p>Follow-up, specifying, opinion/value,</p>	<p>To elicit a range of extrinsic mediating factors perceived by the learners</p>	<p>2</p>

b) What do you think of the role culture, such as the exam-oriented culture and the Chinese culture, plays? 你認為文化上，例如香港的應試文化和儒家文化，又是否對你的準備工作帶來影響？	direct		
--	--------	--	--

The Closing of the Interview 結束	Purpose
<p>Thank you very much for all your contribution. In our discussion of the causes of influences, we have mentioned (a summary of the factors identified). Is there anything that you would like to add or to talk about?</p> <p>非常感謝你們的討論和付出。在剛才關於影響你的備考工作的因素討論中，我們提及到...(總結提及到的因素)。除此以外，你們還有沒有什麼想補充？</p> <p>Thank you. I have no further questions. Do you have any questions before we end the interview? (Pause). If not, let me conclude by telling you how I intend to use your contribution. Following this interview, I will write up a questionnaire based on the influences and the causes of influences you identify. This questionnaire will then be administered to 150 local secondary school students. The results will tell us 1) the influences that the new graded approach in HKDSE-English has on Hong Kong secondary school students, and 2) the causes of such influences. In other words, this project will be a very meaningful one.</p> <p>感謝你們。我準備的問題已全部問完。在我們結束前，你們有無問題想問？(停頓)。如果沒有，我總結一下我將如何運用今日訪問所得的信息。在訪問之後，我將會基於你們所揭示的影響和原因(因素)設計一份問卷。這份問卷將會向香港本地 150 位中學生發放並邀請完成。而問卷的結果將會告訴我們 1) HKDSE 英文科採用的 graded approach 為香港中學生所帶來的影響，2) 這些影響的來源或原因。換言之，這個研究項目將會是十分有意義的。</p> <p>Lastly, please allow me to remind you once again that your identity will be kept strictly anonymous. You also have the right to withdraw from this project at any time. Thank you once again for your contribution. If you have any further questions, you are most welcome me to contact me on my mobile 64929678 or at my e-mail address c.tsang.16@ucl.ac.uk. Thank you and goodbye.</p> <p>最後，容我再提醒一句，你們的資料將會被完全保密，即是匿名。你亦有權在這個項目的任何時間退出這項研究。再次感謝你們的參與。如果你還有什麼疑問，歡迎你們聯絡我，我的電話號碼是 64929678 或者通過電郵 c.tsang.16@ucl.ac.uk 聯絡我。多謝！再見！</p>	<p>To sum up the section by recapitulating the main ideas covered</p> <p>To inform participants the intended use of data</p> <p>To remind participants the guarantee of full anonymity and their right to withdrawal</p> <p>To express gratitude</p>

Appendix 4 – Washback on Student Learning (WSL) Questionnaire

Washback on Student Learning Questionnaire

Background 背景資訊
<p>Since 2012, the graded approach which divides section B of the reading paper and the listening & integrated skills paper into Part B1 (the easier section) and Part B2 (the more difficult section) has been implemented in the Hong Kong Diploma of Secondary Education English Language Examination (HKDSE-English). This 20-minute questionnaire aims at finding out the influences this graded approach has on your learning. There are two parts in this questionnaire. Part A concerns the preparatory work you do in response to the implementation of the graded approach, whereas Part B looks at the factors that motivate you to do so.</p> <p>自 2012 年起，在香港中六文憑試英文科(HKDSE-English)中，閱讀卷(Paper 1)和聆聽及綜合能力考核卷(Paper 3)均採用 graded approach，提供兩份卷別，即 B1(難度較淺)和 B2(難度較深)，給考生選擇其中一份應答。這份問卷旨在探究 graded approach 的實施對你的學習和備考的影響。這份問卷需時 20 分鐘，共有兩大部分：</p> <ul style="list-style-type: none"> • 第一部分 (Part A)：主要關注你為這兩份卷(Paper 1-Reading paper 閱讀卷，和 Paper 3-Listening & integrated skills paper 聆聽及綜合能力考核卷)所做的應考準備工作。 • 第二部分 (Part B)：探究影響你學習和應試準備的因素。 <p>There is no right or wrong answer; all you need to do is to select <u>one</u> option in each item which is the closest to your opinions. After completing the questionnaire, please check to ensure there are <u>no</u> missing items. The information you kindly offer us will have <u>no</u> negative influences on you because your identity is to be kept strictly confidential. You also reserve the perfect right to withdraw from the study at any stage and time.</p> <p>這份問卷沒有任何標準答案。請在每題中選擇一個最貼近你(最能代表你)的選項，並於完成問卷後複檢以確保沒有遺漏。閣下擁有充分的權利在任何時候決定退出這項研究，更不會因此引致任何不良後果。凡有關閣下的資料，將會保密。</p>

Part A – Your Preparatory Work 第一部分–你的準備工作

<p><i>Instructions:</i> Below are statements about the preparatory work you do in response to the implementation of the graded approach. For each item, indicate your extent of agreement with the statement. Please circle the number as appropriate and use the following response format</p> <p>(1=Strongly Disagree, 2=Disagree, 3=Agree, 4=Strongly Agree)</p> <p>以下是有關你為這兩份卷(Paper 1-閱讀卷，和 Paper 3-聆聽及綜合能力考核卷)所做的準備工作的陳述。對於每一項陳述，請根據下列指示指出圈出你有多大程度同意這項陳述：</p> <p>(1=完全不同意, 2=不同意, 3=同意, 4=完全同意)</p>					
1	<p>I practice my preferred section of the mock papers (e.g. school mock papers, mock papers in textbooks) intensively in school English lessons</p> <p>在學校英文堂裡，我會集中操練模擬卷(如：學校模擬卷，課本的模擬題)中我已選定的卷別(B1 或 B2)。</p>	1	2	3	4

2	In school English lessons, I practice my preferred section of the past HKDSE papers hard 在學校英文堂裡，我會集中操練文憑試往年試卷中我已選定的卷別(B1 或 B2)。	1	2	3	4
3	I learn in the school English lessons only when my preferred section is the focus of the lesson 只有當課堂內容針對我已選定的卷別(B1 或 B2)時，我才認為我在該堂學校英文堂裡有所得著。	1	2	3	4
4	When my teacher is commenting on how well I performed in my preferred section, I think I am learning 當老師評鑑我在我已選定的卷別(B1 或 B2)中的表現時，我會認為我在該堂英文堂有所得著。	1	2	3	4
5	I do not listen to my school English teacher when he/she is teaching my dispreferred section 當學校英文老師在課堂裡教授我不選擇/放棄的卷別(B1 或 B2)時，我不會聽書。	1	2	3	4
6	I am reluctant to practice my dispreferred section 我抗拒操練我不選擇/放棄的卷別(B1 或 B2)。	1	2	3	4
7	I do not practice my dispreferred section seriously 我不會認真操練我不選擇/放棄的卷別(B1 或 B2)。	1	2	3	4
8	I drill only on the skills specific to my preferred section (e.g. Reading Part B2 – referencing, inferencing questions and long questions; Listening & Integrated Skills Part B2 – extracting points, analyzing graphs, and summarizing main ideas) 我只操練我已選定的卷別中(B1 或 B2)會出現的題型 (例如：閱讀卷 Part B2 – 推斷型題目，長問題；聆聽及綜合能力考核卷 Part B2 – 提取信息，分析圖表及總結中心思想)。	1	2	3	4
9	I enroll in tutorial classes that focuses on my preferred section 我會報名參加針對我已選定的卷別(B1 或 B2)的補習班。	1	2	3	4
10	I continue to enroll in a tutorial class when the tutor teaches my preferred section well 當補習老師能充分教授我已選擇的卷別(B1 或 B2)時，我會繼續報讀該補習班。	1	2	3	4
11	I do further exercises that are beneficial to my preferred section even when I am out-of-class 即使在課餘時間，我仍會做對我選定的卷別(B1 或 B2)有利的練習。	1	2	3	4
12	I listen to radio programs to prepare myself for my preferred section in the listening & integrated skills paper 我會通過收聽電台節目來準備我在聆聽及綜合能力考核卷(Paper 3)中已選定的卷別(B1 或 B2)。	1	2	3	4

13	I read newspaper articles to prepare myself for my preferred section in the reading and listening & integrated skills paper 我會通過閱讀報刊文章來準備我在閱讀卷(Paper 1)和聆聽及綜合能力考核卷(Paper 3)中已選定的卷別(B1 或 B2)。	1	2	3	4
14	I prepare a vocabulary book to prepare myself for my preferred section in the reading and listening & integrated skills paper 我會通過建立詞彙簿，收集生字來準備我在閱讀卷(Paper 1)和聆聽及綜合能力考核卷(Paper 3)中已選定的卷別(B1 或 B2)。	1	2	3	4
15	I memorize words from the dictionary to prepare myself for my preferred section in the reading and listening & integrated skills paper 我會通過背誦字典中的生字來準備我在閱讀卷(Paper 1)和聆聽及綜合能力考核卷(Paper 3)中已選定的卷別(B1 或 B2)。	1	2	3	4

Part B – Factors Influencing Your Preparatory Work 第二部分-影響你學習和應試準備的因素

Instructions: Below are factors that might influence what you do to prepare yourself for the graded approach in HKDSE-English. For each item, indicate the extent to which you think the factor influences your preparatory work. Please circle the number as appropriate and use the following response format

(1=Not influential, 2=Not quite influential, 3=Quite influential, 4=Very influential)

以下是有可能影響你對 Part B1 或 Part B2 所做的準備工作的因素。對於每一項陳述，請根據下列指示指出(圈出)你有多大程度同意這項因素影響你的應考準備工作：

(1=完全無關, 2=不太影響, 3=有一定的影響, 4=很大影響)

1	My personality 我的性格	1	2	3	4
2	My ambitions 我的目標	1	2	3	4
3	My priority of the English subject (with respect to other subjects) 英文科在我心目中的位置(對比其他科目)	1	2	3	4
4	My knowledge of the two examination sections (i.e. Part B1 and Part B2) in HKDSE-English 我對文憑試英文科中這兩卷別(即 B1 和 B2)的了解	1	2	3	4
5	My perception of the two examination sections (i.e. Part B1 and Part B2) in HKDSE-English 文憑試英文科中這兩卷別(即 B1 和 B2)在我心目中的形象	1	2	3	4
6	My performance in classroom assessment tasks 我在課堂評估中的表現	1	2	3	4
7	My performance in school mock examination 我在學校模擬試中的表現	1	2	3	4
8	My English language ability 我的英文水平	1	2	3	4

9	My study load (availability of time for the English subject) 我的學習負擔 (例如: 可投放在英文科的時間)	1	2	3	4
10	My interest in the English subject 我對英文科的興趣	1	2	3	4
11	Classmates' performance in school examination 同學在學校考試中的表現	1	2	3	4
12	Classmates' selection of examination section (i.e. Part B1 or Part B2) 同學對考試卷別的選擇 (即 B1 或 B2 的選定)	1	2	3	4
13	Classmates' preparatory work 同學的應考準備工作	1	2	3	4
14	Teacher's preference for examination section (i.e. Part B1 or Part B2) 老師對卷別選擇的取向	1	2	3	4
15	Teacher's expectations 老師的期望	1	2	3	4
16	Teacher's assessment of my English language ability 老師對我英文水平的評估	1	2	3	4
17	Teacher's advice 老師的意見	1	2	3	4
18	Comparisons made by the teacher (e.g. with previous years; between classmates) 老師所作的比較(例如: 與往屆學生的比較, 同學間的比較)	1	2	3	4
19	Materials used in class 課堂使用的教材	1	2	3	4
20	School banding 學校排名	1	2	3	4
21	Advertisements posted by tutorial schools 補習社的廣告	1	2	3	4
22	Tutorial schools' advice 補習社的建議	1	2	3	4
23	Skills learnt in tutorial classes 在補習社學到的技巧	1	2	3	4
24	Parents' expectations 家長期望	1	2	3	4
25	Sibling's performance in the HKDSE-English in previous years 兄長在往年文憑試英文科中的表現	1	2	3	4
26	Sibling's advice 兄長的建議	1	2	3	4
27	Posts on forums 網絡論壇上的帖 (e.g. 名校 secrets, 高登)	1	2	3	4
28	Reports published by the media 媒體的報導	1	2	3	4
29	The importance of HKDSE-English in modern Hong Kong society 文憑試英文科在現今香港社會中的重要性	1	2	3	4
30	The examination-oriented culture in Hong Kong 香港注重考試的風氣	1	2	3	4

End of the Questionnaire

Appendix 5 – Approach Letter

16th May, 2017

Dear Principal,

Invitation to Participate in a Research Project on Washback

I am writing to cordially invite your school to participate in my research project entitled “Examining Washback on Learning from a Sociocultural Perspective: The Case of a Graded Approach to English Language Testing in Hong Kong” which is conducted under the supervision of Dr. Talia Isaacs at the Institute of Education, University College London. This research project, which investigates the washback effects (i.e. the influences) brought about the graded approach in the Hong Kong Diploma of Secondary Education English Language Examination (HKDSE-English), will offer the much-needed insights for both the Hong Kong education context and the field of language assessment. Therefore, your participation in this research project will be very much appreciated.

The aim of the research project is to investigate the effects the new graded approach in HKDSE-English has on Hong Kong secondary school learners. In particular, this study attempts to uncover how the graded approach in HKDSE-English influences students’ learning, and what the range of factors shaping such influences are. Yet, to be able to extract the possible washback effects and influencing factors that are representative and generalizable across the wider student population, it is important to include participants from schools with different bandings. Hence, it would be wonderful if I could include students from your school in my pool of participants so that a broader picture can be achieved.

Participation in this research project is voluntary. It will involve, first, three 20-minute semi-structured focus group interviews (N=4 in each group) with a total of 12 Form 4 to 6 students, and, subsequently, the administration of a questionnaire to around 120 students at the school. The study involves no potential risk but in case if the participants feel uncomfortable during the study, they are entitled to stop at any point. As for issues of confidentiality, all data will be kept strictly confidential and stored in the researcher’s encrypted personal computer until 29th August. While the results of the study might be used in future reports and presentations, full anonymity of the participants is assured. More detailed information about the aim, the main research questions, the data collection procedure, and the ethics of the research project can be found in the information sheet attached in this email.

If you would agree to participate in this research project, please complete the attached consent form and return it to c.tsang.16@ucl.ac.uk. Should you have any questions regarding the project, or would like additional information to assist you in reaching a decision about participation, please contact me at +852 64929678 or by e-mail at c.tsang.16@ucl.ac.uk. I very much look forward to your reply and thank you in advance for your assistance in this project.

Kind regards,

Chi Lai, Tsang

Appendix 6 – Information Sheet (for Schools)

Examining Washback on Learning from a Sociocultural Perspective: The Case of a Graded Approach to English Language Testing in Hong Kong

16th May to 29th August

Information sheet for Participating Schools

Who is conducting the research?

My name is Chi Lai Tsang and I am cordially inviting your school to take in part in my research project, 'Examining Washback on Learning from a Sociocultural Perspective: The Case of a Graded Approach to English Language Testing in Hong Kong'. I am currently a MA Applied linguistics student at the Institute of Education, University College London and my research interests fall primarily in the areas of language assessment and language policy and planning.

I am hoping to find out through this research the effects the new graded approach in the Hong Kong Diploma of Secondary Education English Language Examination (HKDSE-English) has on local secondary learners. In particular, this study attempts to uncover how the graded approach influences students' learning, and what the range of factors shaping such influences are.

I very much hope that your school would like to take part. This information sheet will try and answer any questions you might have about the project, but please don't hesitate to contact me if there is anything else you would like to know.

Please explain the research to your students and discuss whether or not they want to take part. I will also ask the students before the task/interview and make it clear that they can drop out if they wish with no negative consequences.

Why are we doing this research?

The washback effects that the graded approach in HKDSE-English has on Hong Kong secondary school's students' learning has yet to be investigated. Therefore, the present mixed methods study is a modest attempt to address this gap. The two main research questions which guide the present study are "What are the washback effects that Hong Kong secondary school learners perceive to have following the introduction of the graded approach in HKDSE-English?" and "What is the full range of intrinsic and extrinsic mediating factors that shape such self-perceived washback effects?". Through carefully and empirically investigating the issue from a sociocultural perspective, this study aims at enriching the understanding of both the washback effects brought about by the new graded approach, and the washback effects on learning in general.

Why am I being invited to take part?

Your participation will contribute significantly to the present study. As the study employs maximum-variation sampling, ideally secondary schools from all three bandings will be included. Therefore, your participation will offer precious data and provide a broader picture of how Hong Kong students are influenced by the graded approach which is indeed much-needed in the field.

What will happen if I choose to take part?

There will be two phases of data collection at the school. In Phase 1, three 20-minute semi-structured focus group interviews (N=4 in each group) with a total of 12 Form 4 to 6 students will be conducted. The aim of these interviews is to elicit a wide range of 1) students' self-perceived influences caused by the graded approach, and 2) factors which shape such influences. Examples of interview questions are "Do you learn differently after the implementation of the graded approach?" and "What are the factors that influence your actions?". Following the identification of the possible influences and underlying factors, a questionnaire will be formulated accordingly. This questionnaire will then be administered to around 120 students at the school in Phase 2 of the study. The aim of this is to screen for the influences and factors that are less relevant in the wider population.

Will anyone know I have been involved?

The identities of the participants are to be kept strictly confidential during the course of the study.

Could there be problems for me if I take part?

The study involves no potential risk but in case if the participants feel uncomfortable during the study, they are entitled to stop at any point.

What will happen to the results of the research?

The results of the study might be used in future reports and presentations. Yet, full anonymity, which is achieved by the use of pseudonyms, is assured. The results of the research will be shared with participants upon their wishes. All the raw data will be stored in the researcher's encrypted personal computer and retained until 29th August. The researcher is the only person who has the access to the data.

Do I have to take part?

It is entirely up to you whether or not you choose to take part. We hope that if you do choose to be involved then you will find it a valuable experience. However, if you choose not to take part, there will be no negative repercussions.

Thank you very much for taking the time to read this information sheet.

If you would like to be involved, please complete the following consent form and return to c.tsang.16@ucl.ac.uk by 16th June, 2017.

If you have any further questions before you decide whether to take part, you can reach me at c.tsang.16@ucl.ac.uk (email) or +852 64929678 (mobile).

This project has been reviewed and approved by the UCL IOE Research Ethics Committee [\[insert reference number\]](#).

Appendix 7 – Consent Forms (for School)

Examining Washback on Learning from a Sociocultural Perspective: The Case of a Graded Approach to English Language Testing in Hong Kong

16th May to 29th August

If your school is happy to participate, please complete this consent form and return to Mr. Chi Lai Tsang at c.tsang.16@ucl.ac.uk by 16th June.

	Yes	No
I have read and understood the information sheet about the research	<input type="checkbox"/>	<input type="checkbox"/>
I give permission for my students to be interviewed	<input type="checkbox"/>	<input type="checkbox"/>
I allow the interviews to be audio-recorded	<input type="checkbox"/>	<input type="checkbox"/>
I give permission for my students to complete the questionnaire	<input type="checkbox"/>	<input type="checkbox"/>
I understand that if any of my students' words are used in reports or presentations they will not be attributed to my students	<input type="checkbox"/>	<input type="checkbox"/>
I understand that my students can withdraw from the project at any time, and that if they choose to do this, any data they have contributed will not be used	<input type="checkbox"/>	<input type="checkbox"/>
I understand that I can contact Mr. Chi Lai Tsang at any time	<input type="checkbox"/>	<input type="checkbox"/>
I have discussed the information sheet with my students	<input type="checkbox"/>	<input type="checkbox"/>

Name _____

Signed _____

Date _____

Researcher's name Mr. Chi Lai Tsang

Signed  _____

Appendix 8 – Information Sheet (for Parents)

Examining Washback on Learning from a Sociocultural Perspective: The Case of a Graded Approach to English Language Testing in Hong Kong

16th May to 29th August

Information sheet for Parents

Who is running this project?

My name is Chi Lai Tsang and I am cordially inviting your child to take in part in my project, 'Examining Washback on Learning from a Sociocultural Perspective: The Case of a Graded Approach to English Language Testing in Hong Kong'. I am currently a MA Applied linguistics student at the Institute of Education, University College London and my research interests fall primarily in the areas of language assessment and language policy and planning.

I am hoping to find out through this project the effects the new graded approach in the Hong Kong Diploma of Secondary Education English Language Examination (HKDSE-English) has on local secondary learners. In particular, this project attempts to uncover how the graded approach influences students' learning, and what the range of factors shaping such influences are.

I very much hope that you would allow your child to take part in this project. This information sheet will try and answer any questions you might have about the project, but please don't hesitate to contact me if there is anything else you would like to know.

Please explain the project to your child and discuss whether or not he/she wants to take part. I will also ask him/her before the task/interview and make it clear that he/she can drop out if he/she wishes with no negative consequences.

Why are we doing this project?

The influences that the graded approach in HKDSE-English has on Hong Kong secondary school's students' learning has yet to be investigated. Therefore, this project is a modest attempt to understand this. The two main questions which guide the project are "What are the influences the graded approach in HKDSE-English has on Hong Kong secondary school learners?" and "What is the range of factors that shape learners' such influences?".

Why is my child being invited to take part?

Your participation will contribute significantly to the project. In particular, your participation will offer precious data and provide a broader picture of how Hong Kong students are influenced by the graded approach which is indeed much-needed in the field.

What will happen if my child chooses to take part?

There will be two phases of data collection at the school. In Phase 1, a 20-minute group interview will be conducted. The aim of the interview is to elicit a wide range of 1) students' self-perceived influences caused by the graded approach, and 2) factors which shape such influences. Examples of interview questions are "Do you learn differently after the implementation of the graded approach?" and "What are the factors that influence your actions?". Following the identification of the possible influences and underlying factors, a questionnaire will be formulated accordingly. This questionnaire will then be administered at the school in Phase 2 of the study. The aim of this is to screen for the influences and factors that are less relevant in the wider population.

Will anyone know I have been involved?

The identity of your child is to be kept strictly confidential during the course of the project.

Could there be problems for me if I take part?

The project involves no potential risk but in case if your child feels uncomfortable during the project, he/she is entitled to stop at any point.

What will happen to the results of the project?

The results of the project might be used in future reports and presentations. Yet, full anonymity, which is achieved by the use of pseudonyms, is assured. The results of the project will be shared with s upon their wishes. All the raw data will be stored in the researcher's encrypted personal computer and retained until 29th August. The researcher is the only person who has the access to the data.

Do I have to take part?

It is entirely up to you whether or not you choose to take part. We hope that if you do choose to be involved then you will find it a valuable experience. However, if you choose not to take part, there will be no negative repercussions.

Thank you very much for taking the time to read this information sheet.

If you would like to be involved, please complete the following consent form and return to c.tsang.16@ucl.ac.uk by 16th June, 2017.

If you have any further questions before you decide whether to take part, you can reach me at c.tsang.16@ucl.ac.uk (email) or +852 64929678 (mobile).

This project has been reviewed and approved by the UCL IOE Research Ethics Committee [insert reference number].

Appendix 9 – Consent Form (for Parents)**Examining Washback on Learning from a Sociocultural Perspective: The Case of a Graded Approach to English Language Testing in Hong Kong**16th May to 29th August

If you allow your child to participate, please complete this consent form and return to Mr. Chi Lai Tsang at c.tsang.16@ucl.ac.uk by 16th June.

	Yes	No
I have read and understood the information sheet about the project	<input type="checkbox"/>	<input type="checkbox"/>
I give permission for my child to be interviewed	<input type="checkbox"/>	<input type="checkbox"/>
I allow the interviews to be audio-recorded	<input type="checkbox"/>	<input type="checkbox"/>
I give permission for my child to complete the questionnaire	<input type="checkbox"/>	<input type="checkbox"/>
I understand that if any of child's words are used in reports or presentations they will not be attributed to my child	<input type="checkbox"/>	<input type="checkbox"/>
I understand that my child can withdraw from the project at any time, and that if he/she chooses to do this, any data he/she has contributed will not be used	<input type="checkbox"/>	<input type="checkbox"/>
I understand that I can contact Mr. Chi Lai Tsang at any time	<input type="checkbox"/>	<input type="checkbox"/>
I have discussed the information sheet with my child	<input type="checkbox"/>	<input type="checkbox"/>

 Name _____

Signed _____

Date _____

 Researcher's name Mr. Chi Lai Tsang

Signed  _____